**UNITED STATES ENVIRONMENTAL PROTECTION AGENCY**
**WASHINGTON, D.C. 20460**

**OFFICE OF**
**CHEMICAL SAFETY AND**
**POLLUTION PREVENTION**

**MEMORANDUM**

**Date: September 15, 2015**

**SUBJECT: Literature Review on Neurodevelopment Effects & FQPA Safety Factor Determination for the Organophosphate Pesticides**

**PC Code:** See Below
**Decision No.:**
**Petition No.:** None
**Risk Assessment Type:** None
**TXR No.:** None
**MRID No.:** None

**DP Barcode:** 331251
**Registration No.:** None
**Regulatory Action:** None
**Case No.:** None
**CAS No.:** See Below
**40 CFR:** None

**FROM:** Anna Lowit, Ph.D., Senior Scientist
Khin Swe Oo, MD, DABT, Toxicologist
Elizabeth Holman, DrPH, Physical Scientist
Health Effects Division (7509P)
Office of Pesticide Programs

Virginia C Moser, PhD, DABT, Fellow ATS, Toxicologist
Toxicity Assessment Division
Office of Research and Development

**THROUGH:** Dana Vogel, Director
Health Effects Division (7509P)

**TO:** Kelly Ballard, Chemical Review Manager
Neil Anderson, Branch Chief
Risk Management and Implementation Branch 1
Pesticide Re-Evaluation Division (7508P)

This paper supports the use of the 10X FQPA Safety Factor in the individual organophosphate human health risk assessments.

| Chemical | PC Code | CAS No. |
|---|---|---|
| **Dicrotophos** | 035201 | 141-66-2 |
| **Fosthiazate** | 129022 | 98886-44-3 |
| **Coumaphos** | 036501 | 56-72-4 |
| **Terbufos** | 105001 | 13071-79-9 |
| **Profenofos** | 111401 | 41198-08-7 |
| **Bensulide** | 009801 | 741-58-2 |
| **Diazinon** | 057801 | 333-41-5 |
| **Ethoprop** | 041101 | 13194-48-4 |
| **Dimethoate** | 035001 | 60-51-5 |
| **Malathion** | 057701 | 121-75-5 |
| **Phosmet** | 059201 | 732-11-6 |
| **Chlorethoxyfos** | 129006 | 54593-83-8 |
| **Acephate/ Methamidiphos** | 103301/ 101201 | 30560-19-1/ 10265-92-6 |
| **Pirimiphos-methyl** | 108102 | 29232-93-7 |
| **TCVP** | 083701 | 961-11-5 |
| **Tribufos** | 074801 | 78-48-8 |
| **Phorate** | 057201 | 298-02-2 |
| **Phostebupirim** | 129086 | 96182-53-5 |
| **DDVP** | 084001 | 62-73-7 |
| **Naled** | 034401 | 300-76-5 |
| **Trichlorfon** | 057901 | 52-68-6 |
| **Fenamiphos** | 100601 | 22224-92-6 |
| **AZM** | 058001 | 86-50-0 |
| **Methidathion** | 100301 | 950-37-8 |
| **Propetamphos** | 113601 | 31218-83-4 |
| **ODM** | 058702 | 301-12-2 |
| **Disulfoton** | 032501 | 298-04-4 |
| **Methyl parathion** | 053501 | 298-00-0 |
| **Temephos** | 059001 | 3383-96-8 |
| **Chlorpyrifos-methyl** | 059102 | 5598-13-0 |

September 15, 2015
D331251

# Literature Review on Neurodevelopment Effects & FQPA Safety Factor Determination for the Organophosphate Pesticides

Office of Pesticide Programs
US Environmental Protection Agency

**Table of Contents**

# Literature Review on Neurodevelopment Effects & FQPA Safety Factor Determination for the Organophosphate Pesticides

## 1.0 Introduction and Background

Organophosphate pesticides (OPs), widely used in agricultural and household pesticidal applications, act by inhibiting acetylcholinesterase (AChE) in nerve cells. Historically the agency has used inhibition of AChE as the point of departure for OP human health risk assessments (HHRAs). This science policy is based on decades of work which shows that AChE inhibition is the initial event in the pathway to acute cholinergic neurotoxicity. AChE inhibition is most often used as the regulatory endpoint for deriving points of departure (PODs) for the single chemical OP HHRAs. In addition, because OPs share the ability to inhibit AChE via phosphorylation of the active site of the enzyme leading to accumulation of acetylcholine and ultimately neurotoxicity, this class of pesticides is subject to assessment of cumulative risk (USEPA, 1999; 2006).

Newer lines of research on OPs in the areas of potential modes of action/adverse outcome pathways (MOAs/AOPs),[1] *in vivo* animal studies, and notably epidemiological studies in mothers and children, have raised uncertainty about the agency's risk assessment approach with regard to the potential for neurodevelopmental effects in fetuses and children. Many of these studies focus on chlorpyrifos and have been the subject of review by the agency over the last several years.

Specific to chlorpyrifos, the agency has taken a stepwise, objective and transparent approach in evaluating, interpreting, and characterizing the strengths and uncertainties associated with all of the available lines of scientific information related to the potential for adverse neurodevelopmental effects in infants and children. The stepwise evaluation began with the September 2008 FIFRA Scientific Advisory Panel (SAP) meeting involving a preliminary review of the literature for chlorpyrifos, with a particular focus on women and children (USEPA, 2008), followed by the draft "Framework for Incorporating Human Epidemiologic & Incident Data in Health Risk Assessment" for integration of epidemiology with other types of experimental data (USEPA, 2010; FIFRA SAP 2010a,b). After the draft framework (2010) was published, the agency released "Chlorpyrifos: Preliminary Human Health Risk Assessment for Registration Review," focusing on the AChE inhibiting potential of chlorpyrifos (USEPA, 2011). This focus was consistent with the recommendation from the 2008 SAP that AChE data provide the most appropriate endpoint and dose-response data for deriving PODs for purposes of risk assessment. In 2012, the agency convened another meeting of the FIFRA SAP focused on chlorpyrifos which incorporated the newest experimental data related to AChE inhibition and both cholinergic and non-cholinergic adverse outcomes, including neurodevelopmental studies

[1] Mode of action (MOA) and adverse outcome pathways (AOPs) describe a set of measureable key events that make up the biological processes leading to an adverse outcome and the causal linkages between such events.

on behavior and cognition effects (FIFRA SAP 2012). Similarly, the agency also performed a more in-depth analysis of the biomonitoring data and of epidemiological studies from three major children's health epidemiology cohort studies in the U.S., as well as plausible hypotheses on MOAs/AOPs leading to neurodevelopmental outcomes (USEPA 2012). Following the 2012 SAP meeting, the agency solicited additional input from federal experts in the areas of Magnetic Resonance Imaging (MRI) and neurobehavioral testing in children to further clarify results obtained by examination of the epidemiological studies.[2] In December, 2014, the agency released "Chlorpyrifos: Revised Human Health Risk Assessment for Registration Review" which included the use of a physiologically-based pharmacokinetic/pharmacodynamic (PBPK-PD) model to derive human PODs, which obviated the need for the animal to human extrapolation factor, and refined intra-species factors for some lifestages (USEPA 2014). The chlorpyrifos 2014 revised HHRA also included retention of the 10X FQPA Safety Factor due to uncertainty regarding the degree of protection the endpoint of AChE inhibition provides for potential neurodevelopmental effects (USEPA, 2014).

A review of the scientific literature on potential MOA/AOP leading to effects on the developing brain was conducted for the 2012 FIFRA SAP meeting (USEPA, 2012) and updated for the 2014 chlorpyrifos revised HHRA (USEPA 2014). In short, multiple biologically plausible hypotheses and pathways are being pursued by researchers including: AChE as a morphogen; cholinergic system; endocannabinoid system; reactive oxygen species; serotonergic system; tubulin, microtubule associated proteins and axonal transport. However, no one pathway has sufficient data to be considered more plausible than the others. Among the available studies, there are effects which are either as sensitive as or more sensitive than AChE inhibition. The fact that there are, however, sparse data to support the *in vitro* to *in vivo* extrapolation, or the extrapolation from biological perturbation to adverse consequence, significantly limits their quantitative use in risk assessment. The SAP concurred with the agency in 2008 and 2012 about the lack of definable key events in a MOA/AOP leading to developmental neurobehavioral effects. Since the 2014 literature review, there are no substantive changes in the ability to define and quantify steps in an MOA/AOP leading from exposure to effects on the developing brain. The lack of an established MOA/AOP makes quantitative use of the epidemiology study in risk assessment challenging, particularly with respect to dose-response, critical duration of exposure, and window(s) of susceptibility. The agency will continue to monitor the scientific literature for studies on the AOP for neurodevelopmental effects but this document does not include an updated literature review on this line of evidence.

This document (Section 2.0) provides the literature review of *in vivo* laboratory animal studies and epidemiology studies for OPs other than chlorpyrifos to support the single chemical HHRAs. It also provides an integrated weight of evidence (WOE) analysis for all the OPs to support retention of the 10X FQPA Safety Factor (Section 3.0).

[2] http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2008-0850-0170

## 2.0 Literature Review

In recent years, the National Academy of Sciences has encouraged the agency to move towards systematic review processes to enhance the transparency of scientific literature reviews that support chemical-specific HHRAs to inform regulatory decision making[3]. The NRC defines systematic review as "a scientific investigation that focuses on a specific question and uses explicit, pre-specified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies".[4] Consistent with NRC's recommendations, EPA's Office of Chemical Safety and Pollution Prevention (OCSPP) is currently developing systematic review policies and procedures. In short, OCSPP employs fit-for-purpose systematic reviews that rely on standard methods for collecting, evaluating and integrating the scientific data supporting the agency's decisions. The literature review described here uses concepts consistent with systematic review such as detailed tracking of search terms and which literature have been included or excluded.

### 2.1 Developmental Neurotoxicity (DNT) Research on OPs other than Chlorpyrifos: Laboratory Animal Studies

The literature on neurobehavioral effects of developmental exposure to chlorpyrifos was summarized and discussed at the 2012 FIFRA SAP. More recent studies were added to this summary for the 2014 revised HHRA. At that time, the conclusions were that the animal studies clearly showed neurobehavioral outcomes following developmental exposure to chlorpyrifos, but there were inconsistencies in the types of effects reported (neurological domain altered, direction of change, gender specificity). Furthermore, the studies were conducted with doses that most likely produced at least some amount of AChE inhibition at some time during the exposure based on results of guideline studies submitted for registration. The impact of these observations has lead the agency to evaluate whether or not these conclusions extend to other OP pesticides. In this review, the studies of a number of OP pesticides are summarized.

The search aimed to focus on rodent studies involving prenatal/perinatal exposure to OPs in which the offspring were evaluated with *in vivo* neurobehavioral tests. The search methods and analytical scope for this analysis are consistent with the chlorpyrifos analysis from the 2012 FIFRA SAP and 2014 HHRA. Preweaning measurements of behavioral development were noted but not compiled, since these could reflect effects of current exposure to the pesticide rather than long-term neuronal changes. Information on AChE inhibition in either fetuses/pups or dams during this exposure period was evaluated where available. Sections 2.1.1-2.1.3 describe the studies from the open scientific literature. Section 2.1.4 summarizes relevant results from the DNT guidelines studies submitted for pesticide registration (US EPA guideline 870.6300 and/or OECD guideline 426).

---

[3] NRC 2011. "Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde" ; NRC 2014. "Review of EPA's Integrated Risk Information System (IRIS) Process"

[4] http://dels.nas.edu/Report/Review-Integrated-Risk/18764

### 2.1.1 Literature Search Strategy & Results

To review and evaluate the developmentally neurotoxic effects of other OPs, a search of the open literature was undertaken. Due to the limited number of studies available, the agency did not limit the search to currently registered OPs. The agency is aware that some OPs listed below are no longer registered for use in the US. In addition, the data evaluation records (DERs) for existing guideline DNT studies were collected from OPP files and summarized. The literature search strategy was developed and conducted by a US EPA reference librarian. Databases searched were PubMed, Web of Science (WoS) and ScienceDirect using key words described below. Duplicates were eliminated after the total database was generated.

1. PubMed (751 results)
(((((organophos* OR Cholinesterase Inhibitors OR acetylcholinesterase inhibitor))) AND ((prenatal OR neonatal OR perinatal OR in utero OR fetal OR foetale OR newborn OR infant* OR postnatal OR gestational OR pregnancy[MeSH Terms]))))) AND ((neurodevelop* OR attention OR birth outcome* OR health outcome* OR cognitive OR cognition OR developmental disability* OR fetal growth OR fetal growth OR foetal development OR fetal development OR sex OR social OR intelligence OR memory OR neurological functioning OR psychomotor OR neurotoxicity OR neurobehavior OR behavior OR learning OR Nervous System[MeSH Terms]) OR Neurotoxicity Syndromes[MeSH Terms]) AND ((((((guinea pigs[MeSH Terms]) OR rabbits[MeSH Terms]) OR mice[MeSH Terms]) OR rats[MeSH Terms]) NOT fishes[MeSH Terms])

2. Web of Science (427 results)
#5 #4 AND #3 AND #2 AND #1
*DocType=All document types; Language=All languages;*
#4 TS=(guinea pig* OR rabbit* OR mice OR mouse OR rat* OR rodent*) NOT TS=(fish*)
*DocType=All document types; Language=All languages;*
#3 TS=(neurodevelop* OR attention OR birth outcome* OR health outcome* OR cognitive OR cognition OR developmental disabilit* OR fetal growth OR fetal growth OR foetal development OR fetal development OR sex OR social OR intelligence OR memory OR neurological functioning OR psychomotor OR neurotoxicity OR neurobehavior OR behavior OR learning OR Nervous System OR Neurotoxicity Syndromes)
*DocType=All document types; Language=All languages;*
#2 TS=(prenatal OR neonatal OR perinatal OR in utero OR fetal OR foetale OR newborn OR infant* OR postnatal OR gestational OR pregnan*)
*DocType=All document types; Language=All languages;*
#1 TS=(organophos* OR Cholinesterase Inhibitors OR acetylcholinesterase inhibitor OR chlorpyrifos)
*DocType=All document types; Language=All languages;*

3. Science Direct (19 results)
(ALL(organophos* OR Cholinesterase Inhibitors OR acetylcholinesterase inhibitor OR chlorpyrifos) and ALL(prenatal OR neonatal OR perinatal OR in utero OR fetal OR foetale OR

newborn OR infant* OR postnatal OR gestational OR pregnan*)) AND (neurodevelop* OR attention OR birth outcome* OR health outcome* OR cognitive OR cognition OR developmental disabilit* OR fetal growth OR fetal growth OR foetal development OR fetal development OR sex OR social OR intelligence OR memory OR neurological functioning OR psychomotor OR neurotoxicity OR neurobehavior OR behavior OR learning OR Nervous System OR Neurotoxicity Syndromes) AND (ALL(guinea pig* OR rabbit* OR mice OR mouse OR rat* OR rodent) and not ALL(fish*)).

This broad literature search identified 1012 potential papers, which were reviewed individually. Specific criteria were applied to select suitable studies, as was previously done with chlorpyrifos. Since the literature on chlorpyrifos has been previously reviewed, those papers were excluded here. This resulted in 19 relevant papers with the following specifications:

- Exposure occurred during gestation and/or the postnatal time frame, ending no later than weaning.
- Dosing included maternal and/or pup administration.
- Dosing was via oral or subcutaneous injection. One paper with intracisternal injection was excluded.
- Behavioral testing of the offspring occurred after weaning and/or into adulthood.
- Studies involved only single-chemical exposure, and where two or more chemicals were administered together, only the single-chemical data were included in the summaries.
- Test subjects were rats or mice. Several papers in pigs and rabbits were excluded due to the lack of comparative database for those species.
- The test measures of interest were neurobehavioral endpoints. At least two studies involved only electrophysiological measures, and those were excluded. No neurochemical, genomic, or other molecular endpoints were included.

The OPs examined, and the number of papers for each, are listed below. Of particular interest are studies from one laboratory (Duke University) that included parathion and diazinon, and can be directly compared to studies with chlorpyrifos using similar experimental designs. The majority of studies used rats (13), and exposures periods varied about evenly between gestational and postnatal stages.

- Parathion (5)
- Diazinon (5)
- Methyl parathion (3)
- Methamidophos (2)
- Chlormephos (1)
- Dichlorvos (1)
- Fenitrothion (sumithion) (1)
- Oxydemeton-methyl (demeton-S-methyl, metasystox-R) (1)

These papers dated back to 1968, and there was a wide range in study quality. Shortcomings were noted in almost all papers, including cursory methodological information and presentation of results, inappropriate statistical analyses, contradictory statements, and problematic interpretation of the data. Regardless, the literature is summarized below in terms of the functional domains organized by each neurobehavioral evaluation.

Below are study descriptions and summaries organized by neurological domain. Appendix 1 is an overall summary for each chemical, presenting each endpoint and outcome.

### 2.1.2 Integration of Literature: Neurobehavioral Domains

*Cognition*

Fifteen studies measured some aspect of cognition: tests included mazes (radial arm maze, Lashley maze, T-maze spontaneous alternation, M-maze), conditioned response (passive avoidance, conditioned avoidance, operant responding, T-maze), and recognition (novel object). Most of these showed adverse effects of OP exposure, although not always in a consistent or dose-responsive manner.

Rats treated with diazinon (0.5, 2 mg/kg/d, postnatal day (PND) 1-4) showed no differences alternating in a T-maze (Timofeeva *et al.*, 2008a). The same rats were tested several months later in a radial arm maze, and showed increased working memory errors (both males and females), but only at the low dose (0.5 mg/kg/d) with no effect on reference memory performance. Diazinon (1 mg/kg/d) was given to rats on gestational day (GD) 15-18 or PND1-4, and there was no change in the trials to criterion in a passive avoidance test; however, there was clearly decreased step-down latency when tested 24 hr later (Vatanparast *et al.*, 2013). This finding suggests a change in memory but not learning. After *in utero* exposure this effect was only seen in females. In contrast, both genders (greater effect in males) were affected following postnatal exposure. Using a novel object test, male mice (females not tested) exposed postnatally (PND8-11) to diazinon (0.5, 5 mg/kg/d) showed less exploration and discrimination of the new object (Win-Shwe *et al.*, 2013). This was significant at both doses (but no dose-response) when tested at PND49, and only the high dose group showed effects at PND84. Mice exposed to diazinon (0.18, 9 mg/kg/d) throughout gestation were tested in a Lashley III maze, with no changes in the number of errors, suggesting no effect on learning (Spyker and Avery, 1977). Thus, these data on diazinon suggest an effect on learning and/or memory (radial arm maze, passive avoidance, novel object recognition) but no changes in learning a maze task. There was a lack of dose-response across studies.

Parathion exposure in rats (0.1, 0.2 mg/kg/d, PND1-4) produced no changes in T-maze spontaneous alternation (Timofeeva *et al.*, 2008b). These rats showed decreased working memory errors, indicative of improvement, at the low dose only (both males and females), and no changes in reference memory errors when tested at about 3 months of age. However, Levin *et al*. (2008) tested littermates from the Timofeeva study beginning at 14 months, and reported increased working memory errors in male rats treated with the low dose only. Interestingly, there is a difference in direction of change and gender specificity compared to the Timofeeva data. Reference memory errors were also increased at both doses (but no dose-response) in male rats only. When the rats were tested again at 17 months, working memory errors were increased at both doses (but no dose-response), again only in males. There was no effect on reference errors at 17 months, and no change in either parameter when the rats were again

tested at 19 months. In a study by Stamper *et al*. (1988), rats (only males were tested) exposed postnatally to parathion (1.3, 1.9 mg/kg/d, PND5-20) showed decreased spontaneous alternations in a T-maze at both doses (dose-response evident), and increased working memory errors in a radial arm maze at both doses (but no dose-response). Reference memory errors were not altered. Al-Hachim and Fink (1968) administered parathion (3 mg/kg/d) during either the first, second or third week of gestation in mice, and reported no effect on conditioned avoidance (sex not mentioned) following any exposure period. Overall, the radial arm maze showed effects of parathion in several studies, but the direction of change, specificity of errors, and gender selectivity differed. The results with the T-maze were contradictory, with one study out of two reporting effects.

Radial arm maze performance was affected by methyl parathion given directly to pups PND1-21 using an incrementally increasing dose schedule (Johnson *et al*., 2009). The middle (0.2 to 0.6 mg/kg/d) and high (0.3-0.9 mg/kg/d) dose groups increased both working and reference memory errors. The lowest dose group (0.2 mg/kg/d throughout dosing) also increased reference memory errors; however, on this measure all dose groups had similar averages. Only males were affected on all measures. Rats dosed with methyl parathion (1 mg/kg/d, GD7-15) were trained to go to a specific side in T-maze, and the correct side reversed five times (Crowder *et al*., 1980). Treated rats had more trials to criterion only on second and fifth switch (note, the text claimed there was an effect on the 4$^{th}$ switch, but the figure does not show it as significant). Effects on only certain reversals is difficult to interpret, and may be a reflection of the multiple t-tests used to analyze the data. Males and females were tested but data were not provided for each sex separately. There were no effects on passive or active avoidance in rats (no mention of gender) exposed to methyl parathion (1, 1.5 mg/kg/d, GD6-20) (Gupta *et al*., 1985). The low-dose group only showed slower latency to bar press during operant shaping, and more days to asymptote; however, details of operant training and schedule were not provided, the sample size was extremely small (n=4/group), and high variability was mentioned. Thus, the most consistent effect was seen in the radial arm maze (even though there was no dose-response), and other tests were either negative or the data were inconclusive.

Several other OPs were tested in different cognitive tasks, but there is no more than one report for any specific pesticide. After single-trial passive avoidance training, male rats (females not tested) treated with dichlorvos (8 mg/kg/d, GD6-15) showed faster latency to cross when tested 7 days later, suggesting delayed retention (Lazarini *et al*., 2004). With fenitrothion, male rats (females not tested) exposed gestationally (5, 10, 15 mg/kg/d, GD7-15) were conditioned to climb a pole to escape shock (Lehotzky *et al*., 1989). The mid and high dose groups (dose-response evident) showed more escapes, were faster, and reached criterion faster than controls. This same pattern seen with reacquisition after a period of extinction, during which time there were no group differences. There was no effect of oxydemeton methyl (0.5, 1.5, 4.5 mg/kg/d, GD6-15) on M-maze learning or memory in rats (Clemens *et al*., 1990). There was also no effect on retention of single trial passive avoidance in mice exposed to methamidophos (1 mg/kg/d, PND3-9); however, the retention trial occurred at 3 hr instead of the more standard 24 hr or greater (Lima *et al*., 2013).

**Table 2.1.1 Summary of Cognitive Outcomes**

| | Early gestation ~GD1-7, ~GD8-14, GD6-15 GD7-15 | Late gestation ~GD15-21, GD6-20, GD15-18 | Gestation GD1-birth | Early postnatal PND1-4, PND3-9 | Late postnatal PND8-11, PND5-20 | Postnatal PND1-21 |
|---|---|---|---|---|---|---|
| Radial arm maze | | | | Diazinon: cognitive deficit – rat, low dose only, M & F[1]<br><br>Parathion: improved cognition – rat, low dose only, M & F[2]<br><br>Parathion: cognitive deficit – rat, M not F[3] | Parathion: cognitive deficit – rat, no dose-response, only M tested[4] | Methyl parathion: cognitive deficit – rat, dose-response, M only[5] |
| T-maze spontaneous alternation | | | | Diazinon: no effect – rat, M & F[1]<br><br>Parathion: no effect – rat, M & F[2] | Parathion: cognitive deficit – rat, dose-response, only M tested[4] | |
| T-maze learning | Methyl parathion: cognitive deficit – rat, sex not specified[6] | | | | | |
| Lashley III maze | | | Diazinon: no effect – mouse, sex not specified[7] | | | |
| M-maze | Oxydemeton methyl: no effect – rat, M & F[8] | | | | | |
| Active avoidance | Fenitrothion: improved | Methyl parathion: no | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | cognition – rat, dose-response, only M tested[9]<br><br>Parathion: no effect – mouse, sex not specified[11] | effect – rat, sex not specified[10]<br><br>Parathion: no effect – mouse, sex not specified[11] | | | | |
| Passive avoidance | Dichlorvos: cognitive deficit – rat, only M tested[12] | Diazinon: cognitive deficit – rat, F not M[13]<br><br>Methyl parathion: no effect – rat, sex not specified[10] | | Diazinon: cognitive deficit – rat, M & F[13]<br><br>Methamidophos: no effect – mouse, sex not specified[14] | | |
| Novel object recognition | | | | | Diazinon: cognitive deficit – mouse, no dose-response, only M tested[15] | |
| Operant responding | | Methyl parathion: cognitive deficit – rat, sex not specified[10] | | | | |

[1] Timofeeva *et al.*, 2008a
[2] Timofeeva *et al.*, 2008b
[3] Levin *et al.*, 2008
[4] Stamper *et al.*, 1988
[5] Johnson *et al.*, 2009
[6] Crowder *et al.*, 1980
[7] Spyker and Avery, 1977
[8] Clemens *et al.*, 1990
[9] Lehotzky *et al.*, 1989
[10] Gupta *et al.*, 1985
[11] al-Hachim and Fink, 1968
[12] Lazarini *et al.*, 2004
[13] Vatanparast *et al.*, 2013
[14] Lima *et al.*, 2013
[15] Win-Shwe *et al.*, 2013

*Motor Activity*

Despite being a commonly tested measure in these studies, only a few of the tested OPs produced changes in motor activity. Rats exposed to dichlorvos (8 mg/kg/d, GD6-15) showed decreased open field locomotion at weaning (males not females), and decreased locomotion and increased immobility as adults (age not specified, only males tested) while rearing was not affected at either age (Lazarini *et al.*, 2004). Male rats exposed to fenitrothion (5, 10, 15 mg/kg/d, GD7-15) showed decreased horizontal activity in the high-dose group only on PND104, with an apparent but not significant effect at PND26, but no effect at PND36 (Lehotsky *et al.*, 1989). Open field testing of rats exposed to methyl parathion (1 mg/kg/d, GD7-15) showed what appeared to be increases only on PND23 and 54, with no differences on PND30, 44, 65 (also not PND18); however, the data are not compelling since statistics are not provided, and the text refers to the data as "a possible change" (Crowder *et al.*, 1980). Methyl parathion-treated rats (1 mg/kg/d, GD6-20) showed a decrease in locomotor activity "accommodation" (apparently the period that is 15-30 min into the activity session) in the low dose group only (Gupta *et al.*, 1985).

Several OPs consistently produced no changes on motor activity, regardless of exposure or test species. No effects were seen with diazinon in rats exposed gestationally (1 mg/kg/d, GD15-18, Vatanparast *et al.*, 2013) or postnatally (0.5, 2 mg/kg/d, PND1-4, Timofeeva *et al.*, 2008a, or 1 mg/kg/d, PND1-4, Vatanparast *et al.*, 2013), or in mice exposed gestationally (0.18, 9 mg/kg/d, GD1-birth, Spyker and Avery, 1977). As with diazinon, there were no motor activity changes following parathion exposure postnatally in rats (0.1, 0.2 mg/kg/d, PND1-4, Timofeeva *et al.*, 2008b, or 1.3, 1.9 mg/kg/d, PND5-20, Stamper *et al.*, 1988) or in mice treated during either the first, second, or third week of gestation (3 mg/kg/d, Al-Hachim and Fink, 1968). There was no effect on measures of activity reported in rats treated with methamidophos (1 mg/kg/d, GD6-15), although high variability of the measures was discussed (deCastro *et al.*, 2000). Methamidophos (1 mg/kg/d, PND3-9) also produced no effect in mice (Lima *et al.*, 2013). Finally, there was no effect of oxydemeton methyl exposure (0.5, 1.5, 4.5 mg/kg/d, GD6-15) on open field activity (Clemens *et al.*, 1990).

In this review, effects on activity are only considered in tests designed specifically for that purpose. During the course of other behavioral tests, e.g., radial arm or T-maze, speed or latency is often measured. These ancillary activity measures were sometimes altered by treatment, but are not included in this domain, since they are not designed to specifically target motor activity.

**Table 2.1.2. Summary of Motor Activity Outcomes**

| Task/Test Apparatus | Early gestation ~GD1-7, ~GD8-14, GD6-15 GD7-15 | Late gestation ~GD15-21, GD6-20, GD15-18 | Gestation GD1-birth | Early postnatal PND1-4, PND3-9 | Late postnatal PND8-11, PND5-20 |
|---|---|---|---|---|---|
| Open field | Dichlorvos: decreased activity – rat, only M tested[1]<br><br>Fenitrothion: decreased activity – rat, dose-response, only M tested[2]<br><br>Methamidophos: no effect – rat, sex not specified[3]<br><br>Methyl parathion: increased activity – rat, sex not specified[4]<br><br>Oxydemeton methyl: no effect – rat, M & F[5]<br><br>Parathion: no effect – mouse, sex not specified[6] | Diazinon: no effect – rat, M & F[7]<br><br>Parathion: no effect – mouse, sex not specified[6] | Diazinon: no effect – mouse, sex not specified[8] | Diazinon: no effect – rat, M & F[7]<br><br>Methamidophos: no effect – mouse, sex not specified[9] | Parathion: no effect – rat, only M tested[10] |
| Figure-Eight | | | | Diazinon: no effect – rat, M & F[11]<br><br>Parathion: no effect – rat, M & F[12] | |
| Donut | | Methyl parathion: decreased activity – rat, no dose-response, sex not specified[13] | | | |

[1] Lazarini *et al*., 2004
[2] Lehotzky *et al*., 1989
[3] deCastro *et al*., 2000
[4] Crowder *et al*., 1980
[5] Clemens *et al*., 1990
[6] al-Hachim and Fink, 1968
[7] Vatanparast *et al*., 2013
[8] Spyker and Avery, 1977
[9] Lima *et al*., 2013
[10] Stamper *et al*., 1988
[11] Timofeeva *et al*., 2008a
[12] Timofeeva *et al*., 2008b
[13] Gupta *et al*., 1985

*Anxiety/Depression*

In a series of tests in rats, Roegge *et al*. (2008) showed that early postnatal exposure to diazinon (0.5, 2 mg/kg/d, PND1-4) produced behaviors suggesting higher anxiety at the high dose (decreased open arm time in an elevated plus maze), lower fearfulness (decreased time to start eating in novel environment) in both dose groups, and anhedonia (decreased chocolate milk preference) at the low dose only, but not depression (forced swim test). These effects occurred only in males and did not show a clear dose-response for the novelty eating and chocolate milk preference tests. Using the same tests (same laboratory; Timofeeva *et al*., 2008b), parathion exposure (0.1, 0.2 mg/kg/d, PND1-4) in rats (high dose, both sexes) increased time in the open arm in an elevated plus maze (suggesting decreased anxiety); however, there was also an increase in center crossings, indicating hyperactivity, that may confound overall interpretation. These same rats showed no changes in novelty-suppressed feeding or in chocolate milk preference. Thus, in these two studies the pesticides appear to have different effects on this functional domain.

There are only a few reports of these behaviors with other pesticides. Rats treated with methyl parathion (1, 1.5 mg/kg/d, GD6-20) showed faster emergence from a cage, interpreted by the authors as lowered anxiety, in the low dose group only (Gupta *et al*., 1985). With chlormephos exposure (~0.06, 0.6 mg/kg/d in drinking water, one week premating to weaning), adult mice (both males and females) in the high-dose group showed decreased time spent in the open arms and increased time in the closed arms (no change in latency) of an elevated plus maze, suggesting increased anxiety (Ceh *et al*., 2012). Mice exposed to methamidophos (1 mg/kg/d, PND3-9) showed no differences in time spent in either arm of an elevated plus maze, but spent less time in center (Lima *et al*., 2013). This was interpreted by the authors as effect on choosing arms, which they say is a cognitive effect; however, this measure is often interpreted to reflect only activity levels. The same mice showed increased immobility in forced swim test, suggesting depressive-like behaviors. Overall, these results are varied and did not consistently show a dose-response.

**Table 2.1.3. Summary of Anxiety/Depression Outcomes**

| | Late gestation ~GD15-21, GD6-20, GD15-18 | Perinatal premating-weaning | Early postnatal PND1-4, PND3-9 |
|---|---|---|---|
| Elevated plus maze | | Chlormephos: increased anxiety – mouse, M & F[1] | Diazinon: increased anxiety – rat, dose-response, M not F[2]<br><br>Methamidophos: no effect –mouse, sex not specified[3]<br><br>Parathion: decreased anxiety – rat, dose-response, M & F[4] |
| Chocolate milk preference | | | Diazinon: increased anhedonia – rat, no dose-response, M not F[2]<br><br>Parathion: no effect – rat, M & F[4] |
| Novelty suppressed feeding | | | Diazinon: decreased fearfulness – rat, dose-response, M not F[2]<br><br>Parathion: no effect – rat, M & F[4] |
| Forced swim | | | Diazinon: no effect – rat, M & F[2]<br><br>Methamidophos: increased despair – mouse, sex not specified[3] |
| Open field behaviors | Methyl parathion: decreased anxiety – rat, no dose-response, sex not specified[5] | | |

[1] Ceh *et al.*, 2012
[2] Roegge *et al.*, 2008
[3] Lima *et al.*, 2013
[4] Timofeeva *et al.*, 2008b
[5] Gupta *et al.*, 1985

*Social Behavior*

Only one study has used tests of social behavior using these OPs. With exposure to fenitrothion (5, 10, 15 mg/kg/d, GD7-15), rats in mid and high dose (dose-response) spent more time actively interacting in conspecific pairs (Lehotsky *et al*., 1989). The scarcity of data on this measure prevents any conclusions across OPs.

*Sensory Function*

In a study measuring response to a tactile stimulus, with and without an acoustic prepulse, male (but not female) rats treated with diazinon (0.5, 2 mg/kg/d, PND1-4) showed less prepulse inhibition at both doses; however, no dose-response was evident (Timofeeva *et al*., 2008a). Using the same paradigm, rats treated with parathion (0.1, 0.2 mg/kg/d, PND1-4) showed a different pattern: lower response to the stimulus alone (high dose, both sexes), but no change in the inhibition produced by the prepulse (Timofeeva *et al*., 2008b). Mice treated with diazinon through gestation showed no change in response to noise (auditory startle) or smell (olfactory orientation), but did show change in visual cliff behavior (more steps off a "cliff") which occurred only in females in the low dose group (Spyker and Avery, 1977). Additional studies are needed for general conclusions regarding the effects of these pesticides on sensory function.

**Table 2.1.4. Summary of Sensory Outcomes**

| | Gestation<br>GD1-birth | Early postnatal<br>PND1-4,<br>PND3-9 |
|---|---|---|
| Auditory | Diazinon: no effect – mouse, sex not specified[1] | |
| Tactile with or without prepulse | | Diazinon: decreased sensory gating – rat, no dose-response, M not F[2]<br><br>Parathion: decreased startle response – rat, dose-response, M & F[3] |
| Visual | Diazinon: decreased function – mouse, no dose-response, F not M[1] | |
| Olfactory | Diazinon: no effect – mouse, sex not specified[1] | |

[1] Spyker and Avery, 1977
[2] Timofeeva *et al*., 2008a
[3] Timofeeva *et al*., 2008b

*Neuromotor Function*

One study assessed neuromotor function in mice exposed to diazinon (0.18, 9 mg/kg/d, GD1-birth), and reported some changes in motor abilities measured at about 2 months of age; however, these were not consistent in dose or direction of change (Spyker and Avery, 1977). Increased ability was suggested by a longer time to cling to a rod (both doses, no dose-response). In contrast, there was less ability to stay on an increasingly inclined plane (both doses, no dose-response) or perhaps on a rotarod (group means not significant due to large variability, suggestive of effect at both doses).

Rats in the high-dose group exposed to fenitrothion (5, 10, 15 mg/kg/d, GD7-15) fell off a rotarod faster on PND26 and PND104, but not PND36. Only males were tested (Lehotsky *et al.*, 1989).

There were no neuromotor changes in terms of rotarod performance in rats following exposure to parathion (1.3, 1.9 mg/kg/d, PND5-20) (Stamper *et al.*, 1988) or methyl parathion (1, 1.5 mg/kg/d, GD6-20) (Gupta *et al.*, 1985). Overall, there is little support for conclusions of neuromotor outcomes following these pesticides, but more studies are needed.

**Table 2.1.5. Summary of Neuromotor Outcomes**

| | Early gestation ~GD1-7, ~GD8-14, GD6-15 GD7-15 | Late gestation ~GD15-21, GD6-20, GD15-18 | Gestation GD1-birth | Late postnatal PND8-11, PND5-20 |
|---|---|---|---|---|
| Rotarod | Fenitrothion: decreased performance – rat, dose-response, only M tested[1] | Methyl parathion: no effect – rat, sex not specified[2] | Diazinon: no effect – mouse, sex not specified[3] | Parathion: no effect – rat, only M tested[4] |
| Inclined plane | | | Diazinon: decreased performance – mouse, dose-response, sex not specified[3] (Spyker) | |
| Rod cling | | | Diazinon: increased performance – mouse, no dose-response, sex not specified[3] | |

[1] Lehotzky *et al.*, 1989
[2] Gupta *et al.*, 1985
[3] Spyker and Avery, 1977
[4] Stamper *et al.*, 1988

### 2.1.3 Integration with AChE Inhibition

There are some data with which to compare effective doses in these DNT studies with doses producing AChE inhibition. A number of studies in this literature review included AChE inhibition (brain and/or blood) in their measurements. Information on diazinon and parathion can be taken from a separate (non-behavioral) study (Slotkin *et al.*, 2006) conducted in the same laboratory with the same dosing paradigm used in several studies (Levin *et al.*, 2008; Roegge *et al.*, 2008, Timofeeva *et al.*, 2008a, 2008b).

Neurobehavioral effects of diazinon were reported in rats at doses of 0.5-2 mg/kg/d (Roegge *et al.*, 2008; Timofeeva *et al.*, 2008a). Slotkin *et al.* (2006) reported that a dose of 0.5 mg/kg/d produced some (<10%, statistically significant) brain AChE inhibition when measured the day after the last dose. At 2 hr after a higher dose (2 mg/kg/d; lower doses were not tested at 2 hr), there was greater brain inhibition (25-30%) compared to 24 hr (10-20%). Thus it is probable that for diazinon, inhibition during and shortly after the dosing period (i.e., within hours) was greater at lower doses. While there is no direct AChE data following diazinon exposure at 1 mg/kg/d (Vataparast *et al.*, 2013), it can be assumed from the Slotkin data that this dose also inhibited brain AChE at some time during/after dosing. No AChE activity was measured in mice by Spyker and Avery (1977), but the high dose of 9 mg/kg/d resulted in depressed weight gain, a sign of maternal toxicity. There is no mention of toxicity at the lower doses (0.5, 5 mg/kg/d) used in mice by Win-Shwe *et al.* (2013).

The parathion studies showing effects at 0.1 and 0.2 mg/kg/d (Timofeeva *et al.*, 2008b; Levin *et al.*, 2008) are also informed by the AChE inhibition presented in Slotkin *et al.* (2006) in which a dose of 0.1 mg/kg/d inhibited brain AChE 5-15% on the day after the last dose. There are no AChE data available for a higher dose, 0.2 mg/kg/d, but Timofeeva noted 5% mortality in that dose group. Much higher doses (1.3, 1.9 mg/kg/d in rats) were reported to produced 35, 68% brain inhibition on the day after the last dose (PND5-20), with 26, 36% inhibition persisting a week later (Stamper *et al.*, 1988).

The methyl parathion postnatal incrementing dose paradigm in rats used by Johnson *et al.* (2009) produced brain AChE inhibition in all dose groups at the end of dosing, persisting for 10 to 20 days later; recovery was evident 30-40 days later. The low dose (0.2 mg/kg/d throughout) produced 13-15% inhibition. During gestational exposure to higher doses of methyl parathion at 1, 1.5 mg/kg/d, the dams had 20, 60% brain inhibition at the end of dosing (GD6-20) (Gupta *et al.*, 1985). At the high dose, cholinergic signs and increased resorptions were noted. Furthermore, the offspring (fostered to control dams) showed brain inhibition as high as 50% in both dose groups when measured at birth and also PND7, 14, 21, and 28. The low dose showed recovery at PND28, but not the high dose.

In the study of rats treated with fenitrothrion, there was postnatal mortality of 16-17.5% at all doses (5, 10, 15 mg/kg/d, GD7-15), compared to 5% in controls (Lehotsky *et al.*, 1989). In separate study of rats, 5, 25 mg/kg on GD19 produced 40, 80% brain AChE inhibition in dams,

and fetal brains showed about 90% inhibition (no dose response) (Sochaski *et al.*, 2007). Assuming similar responses, the postnatal mortality observed in the Lehotsky study could be at least partly due to this high degree of AChE inhibition in both dams and fetuses.

In rats, nonpregnant females dosed methamidophos 1 mg/kg/d for 10 days showed 16% plasma AChE inhibition, but brain AChE was apparently not measured. This suggests that dams treated at the same dose in a study by deCastro *et al.* (2000) most likely experienced some plasma inhibition. Mouse pups dosed with methamidophos 1 mg/kg/d PND3-9 showed ~36, 46% brain inhibition 1, 4 hr after first dose, 53, 61% inhibition at 1, 4 hr after last dose, and ~19% brain inhibition the day after last dose (Lima *et al.*, 2013). The mice therefore experienced considerable brain AChE inhibition throughout dosing.

In the oxydemeton methyl study, dams showed 22-68% brain AChE inhibition on the day after the last dose (0.5-4.5 mg/kg/d, dose response), and 5 days later (GD20) there was 20-54% brain AChE inhibition. Fetal brains taken the day after dosing showed no inhibition (Clemens *et al.*, 1990); however, there were no fetal AChE tissues collected during or shortly after the dosing period when AChE inhibition would be greatest.

Overall, in the studies for which there are direct or comparable data, it is clear that the dosing paradigms produced AChE inhibition and in some cases maternal toxicity. Indeed, there are no studies reporting or even suggesting a lack of AChE inhibition in the dam and/or fetus/pup at any time during dosing. Thus, it is not known whether exposure paradigms that do not inhibit AChE would produce any neurobehavioral effects.

### 2.1.4 Summary of Findings from the Developmental Neurotoxicity (DNT) Guideline Studies

DNT studies have been submitted for 20 OPs, summarized in Appendix 2. These studies follow the US EPA guideline 870.6300 and/or OECD guideline 426 which require testing of motor activity, acoustic startle response, learning and memory, and brain morphometrics in the offspring around weaning and also in adulthood. In general, these studies provide exposure during development either via diet or oral gavage dosing, including direct dosing of the pup preweaning. As with the literature studies, these submitted studies have shortcomings such as inappropriate statistical analyses, limited methodological information and presentation of results. Many measures tend to show high variability, which reduces their interpretability and utility.

In order to compare the submitted guideline and published studies under the scope identified under Section 2.1 and to be consistent with the chlorpyrifos 2012/2014 review, only changes that occurred after dosing had ended (i.e., shortly after weaning or as adults) were considered here. Across the seven submitted studies that reported effects, there are mostly changes in acoustic startle reactivity, cognitive function, and to a lesser extent, motor activity. Some OPs

altered multiple domains, others only one. There are both submitted guideline studies and literature studies for only four OPs: diazinon, methyl parathion, methamidophos, and dichlorvos. Diazinon produced cognitive changes but no effects on motor activity or acoustic startle in the DNT: these results are generally in agreement with published studies. Male offspring in the high dose group (~33 mg/kg/d in the dam diet) showed increased errors and longer latency in Biel maze performance at both PND24 and PND62, and similar effects were seen in females but only in the middle dose group (~3.4 mg/kg/d via diet) at PND24. Following methamidophos exposure, female rats in the middle and high dose groups (~1.7, 5.2 mg/kg/d in the dam diet) showed decreased peak amplitudes of the startle response, which was statistically significant at PND38 and apparent but not significant at PND60. In contrast to the literature reports of cognitive and motor effects of methyl parathion, there were no reported changes in the submitted guideline study. The submitted study of dichlorvos was uninterpretable due to high pup mortality in all groups, including control.

In these studies, AChE activity was assessed as part of the DNT itself, or by means of a separate study comparing the response in pups and adults (comparative cholinesterase, or CCA, studies). Thus, in almost all guideline DNT studies there are adequate data describing AChE inhibition in the pups at some time during development. It is clear from these DNT studies that the doses used did produce AChE inhibition in the offspring, sometimes at all doses tested. It was noted that in these studies, most of the reported effects occurred before weaning, which is the period during which there was likely to be ongoing AChE inhibition.

Thus, as with the literature studies, there are scant data that could inform potential neurodevelopmental changes occurring at doses lower than those needed to inhibit AChE. Furthermore, there is little consistency in patterns of effects across studies or chemicals. Thus, there is uncertainty as to whether lower, non-inhibiting exposures are developmentally neurotoxic; this uncertainty was described in the chlorpyrifos reviews and remains applicable for the available data for other OPs as well.

### 2.1.5 Conclusions on *In Vivo* Laboratory Animal Studies

For chlorpyrifos, there are >30 papers on developmental neurotoxicity; for the remaining OPs, the literature is sparse with very few studies for each OP (including DNT guideline studies). The studies span over decades, and many of the lower quality studies were the earlier ones; however, some very recent papers also have significant deficits. Methodological detail is lacking, inappropriate statistical analyses are applied, results are cursorily described and/or inaccurately presented, and interpretation of some behavioral changes is faulty. Overall, most studies have significant shortcomings and/or are of low quality.

The most commonly tested behaviors considered aspects of cognition. In the majority of studies, some sort of cognitive deficit was detected, especially with working memory performance (radial arm maze) and conditioned response retention (passive avoidance). However, in many cases there was no dose-response, there was some gender specificity which

did not replicate in multiple studies, and cognitive improvement instead of deficit was noted in a few papers. Changes in motor activity in offspring were generally not reported, and the direction of change differed in the papers reporting such effects. There is generally not enough information to make definitive statements about OP effects on other types of neurological disorders.

Few published papers included AChE measurements of the dams and/or offspring, but where measured, all doses used inhibited AChE to some degree. Some papers even reported overt maternal and fetal toxicity. This was also the case in the guideline studies, most of which included concurrent or supplemental data on AChE inhibition. Since there are no studies with low doses that definitively do not inhibit AChE, there is no information in the animal literature that shows whether or not there would be developmentally neurotoxic outcomes at those lower exposures.

## 2.2 Epidemiology Research on OPs other than Chlorpyrifos

### 2.2.1 Overview of 2012/2014 and 2015 Literature Reviews

In April 2012, EPA presented to the FIFRA Scientific Advisory Panel (SAP) its review and assessment of several epidemiological investigations of the potential adverse neurodevelopmental outcomes of *in utero* and early life exposure to chlorpyrifos. In this effort, EPA limited its review to studies conducted within three major US based prospective birth cohort studies: 1) Mother's and Newborn Study of North Manhattan and South Bronx conducted by Columbia University, referred to in this document as "Columbia Study/Cohort;" 2) Mount Sinai Inner-City Toxicants, Child Growth and Development Study, or the "Mount Sinai Study/Cohort;" and 3) Center for Health Assessment of Mothers and Children of Salinas Valley (CHAMACOS) conducted by the University of California Berkeley, or "CHAMACOS Study/ Cohort." The conclusion of EPA's evaluation, supported by the FIFRA SAP (2008, 2012), was that "chlorpyrifos likely played a role in the neurodevelopmental outcomes observed in these studies." The major findings of these cohorts are briefly summarized in Section 2.2.9 to provide context for integrating the findings of the 2015 review and for the WOE analysis.

In the current review, the agency has expanded its consideration of the epidemiological data to include studies of any OP pesticide; several different types of development and neurological, neurodevelopmental, and neurobehavioral health outcomes; studies performed in non-U.S. countries as well as US based studies; and non-cohort studies.

### 2.2.2 Literature Search Methodology

To identify the epidemiological investigations of the association between OP exposure and adverse neurological, neurodevelopmental or neurobehavioral effects, EPA scientists queried PubMed/Medline and Web of Science directly. In this literature search, emphasis was placed

upon identification of all possible epidemiological studies available, and the ability to use the identical search string in both PubMed/Medline and Web of Science The following search string was utilized:

> ((Chlorpyrifos OR Organophosphates) AND (prenatal OR neonatal OR perinatal OR in utero OR fetal OR foetal OR newborn OR infant* OR infancy OR preschool OR child* OR maternal OR mother* OR pregnan*) AND (neurodevelop* OR attention OR birth outcome* OR health outcome* OR birth height OR birth weight OR birth length OR cephalometry OR head circumference OR child development OR cognitive OR cognition OR developmental disability* OR fetal growth OR foetal growth OR foetal development OR fetal development OR intelligence OR memory OR neurological functioning OR psychomotor) AND (human)) (Filter: English language only)

With the aid of an EPA reference librarian, EPA also searched the following databases: PsycInfo, Agricola, Biosis, Embase, Enviroline, Gale Health and Wellness, Global Health, Pascal and Pollution Abstracts. EPA used similar, but modified search terms as listed above. Upon identification of the final set of relevant articles (n=38), limited hand-searching of the reference lists and citation mapping *(Science citation index)* of articles deemed to be most relevant to the review question was performed.

EPA identified 300 articles across these several biomedical search engines. Removing duplicates, there were 243 articles, and 79 were determined to be epidemiological investigations of potential relevance. The 164 studies excluded from the analysis comprised 57 exposure only studies; 51 review articles; 33 reports of acute OP intoxication; 20 studies in non-human systems; and 3 were otherwise not relevant. Among the 79 potentially relevant epidemiologic studies, 41 were excluded; 17 articles were previously reviewed in 2012; 16 were epidemiological methods papers including exposure validation studies without an original epidemiological risk estimate; and 8 were otherwise not relevant for various reasons. Among the 40 remaining studies, 2 were additionally excluded (one was a duplicate study published a second time; the other did not make a measure of an OP pesticide. Therefore, 38 articles are included in this narrative literature review (referred to herein as "2015 Literature Review/Studies"). The determination of relevance to the study question was made by two EPA epidemiologists who agreed by consensus as to article disposition in the literature search.

The following sections provide the results of this literature review. Section 2.2.3 describes the breadth and depth of the 2015 literature review, with Section 2.2.4 summarizing the approach for assigning a quality ranking and Section 2.2.5 providing the results of this quality ranking. In Section 2.2.6, these studies were further analyzed with focus on identification of the most appropriate exposure assessment and relevant outcomes for this assessment. Studies focusing solely on birth outcomes are discussed in Section 2.2.7. However, the emphasis in this assessment is on those studies focusing on neurodevelopmental outcomes, which are discussed in detail in Section 2.2.8 and summarized in Section 2.2.9.

### 2.2.3 Breadth and Depth of the 2015 Literature Review

Key features of each of the 38 articles in the current 2015 literature review are summarized in Table 2.2.5-1 and 2.2.5-2, as well as Appendix 3. These articles represent 31 distinct studies covering a wide range of study designs, study locations and time periods, and exposure and outcome measurement approaches and are listed in Table 2.2-a. There were 10 different study designs identified; the majority of the studies utilized a cross-sectional or a prospective birth cohort study design.

In addition to four articles on the previously-reviewed studies (two on the Mt. Sinai Cohort and two on the CHAMACOS Cohort), there were a number of new birth cohorts or named studies in this literature, including (see Table 2.2.5-1, 2.2.5-2, and Appendix 3 for details):

- Denmark, Birth Cohort (Andersen *et al*. 2015)
- Saint Peter's University Hospital, New Brunswick, New Jersey, Birth Cohort (Barr *et al*. 2010)
- Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT), Mexico City, Mexico, Birth Cohort (Fortenberry *et al*. 2014, 2014a)
- EcoSalud Project, Cayambe-Tabacundo region, Ecuador, Infant and Young Child Cohort (Handal *et al*. 2007, 2007b, 2008)
- University Hospital of Heraklion, Crete, Greece, Birth Cohort (Koutroulakis *et al*. 2014)
- Children's Pesticide Survey (CPS), Yuma County, Arizona (Lizardi *et al*. 2008)
- Infancio y Medio Ambiente (INMA) (Environment and Childhood), Spain, Birth Cohort (Llop *et al*. 2013)
- Perturbateurs endocriniens: Étude Longitudinale sur les Anomalies de la Grossesse, l'Infertilité et l'Enfance (PELAGIE), Brittany, France, Birth Cohort (Petit *et al*. 2010)
- Health Outcomes and Measure of the Environment (HOME), Cincinnati, OH, Birth Cohort (Rauch *et al*. 2012)
- Embilipitiya Base Hospital, Southern Sri Lanka, Birth Cohort (Samarawickrema *et al*. 2008)
- Ontario Farm Family Health Study, Ontario, Canada, Birth Cohort (Savitz *et al*. 1997)

- Childhood Autism Risks from Genetics and the Environment (CHARGE), California, Child Cohort (Shelton *et al*. 2014)

- Secondary Exposure to Pesticides Among Infants, Children and Adolescents (ESPINA), Pedro Moncayo County, Pichincha, Ecuador, Child Cohort (Suarez-Lopez *et al*. 2012, 2013, 2013a)

- Shenyang, China, Birth Cohort (Zhang *et al*. 2014)

Exposures were assessed via environmental samples, biomarkers, and/or proxy methods.

- Only one study used environmental samples as an exposure measure - methyl parathion in household wipe samples (Ruckart *et al*., 2004).

- Among the studies, 11 biomarkers were used to assess exposure. One study utilized a direct measure of OP pesticide – chlorpyrifos (CPF) in maternal and cord serum (Barr *et al*., 2010). Several studies looked at specific OP parent metabolites (IMPY, 3,5,6-trichloro-2-pyridinol [TCPy], PNP). The majority of studies used non-specific OP biomarkers: dialkyl phosphates (DAPs), diethyl phosphates (DEPs), and dimethyl phosphates (DMPs). Two effect biomarkers, AChE and butyl cholinesterase (BuChE), were used as the exposure metric.

- Proxy exposure methods included questionnaire and non-questionnaire approaches. For example, questionnaire-based exposures included maternal and or paternal self-report of living with an exposed worker, occupational exposure/ employment, and home pesticide use and child outdoor play exposure. Non-questionnaire-based exposures included Community/area of residence, distance to treated area/farm, percent of area treated with pesticide, level of urbanization, pounds OP pesticide used/year, and pesticide spray season.

Likewise, there were numerous outcome measures examined across the studies, falling into six broad categories: birth characteristics, autonomic nervous system (ANS) effects, Attention Deficit Hyperactivity Disorder (ADHD)/attention problems, autism, general neurodevelopment (cognitive, behavioral, IQ), and physiological effects. The most common outcome measures were birth characteristics (with birth weight, birth length, head circumference and gestational age being the most frequent) and neurodevelopment tests and test batteries. Table 2.2-b lists the many specific neurodevelopment tests employed in the studies. Most studies utilized more than one test, and few tests were utilized in more than one study.

Other features that varied widely among the studies include:

- Study periods ranged from 1991 to 2012 (reports were published from 1997 to 2015).

- Study locations included Canada (2 studies), China (3 studies), Costa Rica, Israel, Denmark, Ecuador (3 studies), France, Greece, Mexico (3 studies), Poland, Spain, Sri Lanka, and the United States [AZ, CA (2 studies), MS, NJ, NY, NC, OH (2 studies), OR, National (2 studies)].
- Study sizes varied from 25 to 3,159 participants.
- Children's ages ranged from newborns to age 15 years.

Pre-natal exposures were assessed in 19 reports, post-natal exposures in 13 reports, and both pre- and post-natal exposures in 5 reports.

**Table 2.2-a Study Designs, Exposure Measurement Methods, and Outcome Measurement Methods Used across the 31 2015 Review Studies**

| Study design (# studies used) | Exposure measurement (# studies used) | Outcome measurement (# studies used) |
|---|---|---|
| Prospective Cohort (10)<br>Retrospective Cohort (3)<br>Case-control (2)<br>Cross-sectional (15)<br>Ecological (1) | **Biomarkers**<br>**AChE** - acetyl cholinesterase (maternal 1, child 2)<br>**BuChE** - butyl cholinesterase (child 1)<br>**CPF** -chlorpyrifos parent (maternal, cord serum 1)<br>**TCPy** - chlorpyrifos metabolite (maternal 1, child 1)<br>**DZN** – diazinon parent (1)<br>**IMPY** – diazinon metabolite (child 1)<br>**MAL** - malathion parent (0)<br>**MDA** - malathion metabolite (maternal 1)<br>**DAP** – dialkyl phosphate (maternal 5, child 7, amniotic fluid 1)<br>**DEP** – diethyl phosphate (maternal 5, child 7, amniotic fluid 1)<br>**DMP** – dimethyl phosphate (maternal 5, child 7, amniotic fluid 1)<br>**OP** - organophosphate (cord blood 1)<br>**PNP** - para-nitrophenol (child 2)<br><br>**Environmental**<br>Methyl parathion - Household wipe (1)<br><br>**Proxy (questionnaire)**<br>• Maternal occupational exposure/ employment (6)<br>• Paternal occupational exposure/ employment (2)<br>• Living with exposed worker (1)<br>• Home and outdoor play exposure (1)<br><br>**Proxy (non-questionnaire)** | **Birth characteristics** (10)<br>• Birth Weight/LBW/FGR (records 7, NP 1, report 2)<br>• Birth Length (records 3)<br>• Head circumference (records 3, NP 1)<br>• Abdominal circumference (records 1)<br>• Gestational age (GA)/preterm (report 2, records 3)<br>• Ponderal index (records 1)<br>• Placental maturity index (1)<br>• Spontaneous abortion/miscarriage (report 1)<br>• Altered sex ratio (report 1)<br>**Autonomic Nervous System** (4)<br>**ADHD/attention problems** (2)<br>• Diagnosis DISC-IV (1)<br>• Use of medication (1)<br>• Screening (CPRS-R, CPT, BASC-PRS) (1)<br>**Autism** (2)<br>• CA Department of Developmental Services (CDDS) reports<br>• US Individuals with Disabilities Education Act (IDEA) reports<br>• Autism spectrum disorders (ASD)<br>• Autism Diagnostic Observation Schedule (ADOS) combined with ADI-R.<br>**Neurodevelopmental (ND) test/battery** - US (3) – see Table 2.2-2b for details<br>**ND test/battery - non-US** (11) – see Table 2.2-2b for details<br>**IQ** (4) - see Table 2.2-2b for details |

| | | |
|---|---|---|
| | • Community/area of residence (3)<br>• Distance to treated area/farm (2)<br>• Percent of area treated with pesticide (1)<br>• Level of urbanization (1)<br>• Pounds OP used/year (1)<br>• Pesticide spray season (1) | **Physiological** (2)<br>• AChE activity, child<br>• BuChE activity, maternal<br>• Antioxidant status: superoxide dismutase (SOD) activity<br>• Fetal oxidative stress: malondialdehyde (MDA) concentrations)<br>• Fetal DNA fragmentation: electrophoresis |

## Table 2.2-b Detailed Neurological Outcome Measurement Methods Used across the 31 2015 Review Studies

| Outcome measurement (# studies used) |
|---|
| Neurodevelopmental (ND) test/battery - US (4)<br>• Pediatric Environmental Neurobehavioral Test Battery (PENTB) : cognitive, motor, sensory, and affect domains<br>  ○ Developmental Test of Visual-Motor Integration (VMI)<br>  ○ Kaufman Brief Intelligence test (K-BIT)<br>  ○ Purdue Pegboard<br>  ○ Story Memory and Story Memory-Delay from Wide Range Assessment of Memory and Learning<br>  ○ Trail-Making test, Part A and Part B<br>  ○ Verbal Cancellation test<br>• Children's Memory Scale (CMS)<br>• Behavioral measures:<br>  ○ The Child Behavior Checklist/4-18<br>  ○ The Teacher Report Form<br>• Developmental delay (DD):<br>  ○ Mullen Scales of Early Learning (MSEL)<br>  ○ Vineland Adaptive Behavioral Scale (VABS)<br>  ○ Reciprocal social interaction: Social Responsiveness Scale |
| ND test/battery - non-US (11)<br>• Bayley Scales of Infant Development - mental and psychomotor development<br>• Behavioral Assessment and Research System (BARS): Memory and attention, response speed and coordination, visual memory, attention, divided attention, recall and recognition memory, dexterity, hand-eye coordination (2)<br>• Figure drawing task: child's perception and dexterity<br>• Long-term memory test<br>• Ages and Stages Questionnaire (ASQ) – communication, fine motor, gross motor, problem solving, and personal–social skills<br>• Strengths and Difficulties Questionnaire, parent version (SDQ): behavioral problems<br>• NEPSY-II test (trained examiners): general assessment battery:<br>  ○ attention and executive functioning<br>  ○ language<br>  ○ memory and learning<br>  ○ sensorimotor (visuomotor precision),<br>  ○ visuospatial processing<br>  ○ Statue and Knock Tap<br>• Neonatal Behavioral Neurological Assessment (NBNA)<br>  ○ Behavior<br>  ○ Passive Tone<br>  ○ Active Tone<br>  ○ Primary Reflexes<br>  ○ General Assessment<br>• Reach-and-grasp, bi-manual coordination: Prehension abilities<br>• UC Berkeley Preferential Looking Test Cards: Visual acuity skills<br>• Visual Motor Integration (VMI), Beery-Buktenica, 4th Ed.<br>• Finger Tapping test: Manual motor speed<br>• Catsys equipment: Simple reaction time<br>• Conners' Continuous Performance Test II (CPT II, v5): Attention<br>• Woodcock-Johnson III Tests of Cognitive Abilities (WJ-III) Verbal Comprehension test: Long-term memory and language function<br>• Visuospatial performance and memory functions:<br>  ○ Raven's Colored Progressive Matrices<br>  ○ Stanford–Binet Copying Test, 4th ed<br>• Physical examination (social response, spontaneous motility, involuntary movements, Romberg's sign, walking straight line, standing on one leg, number hops, biceps and patellar reflexes, finger opposition, diadochokinesis, finger–nose coordination, hearing, vision)<br>• Gesell Development Schedule (GDS): motor, adaptive, language, and social<br>• Santa Ana Form Board: Motor coordination<br>• Child's developmental delay - Parent interview |

| IQ (4) |
| --- |
| • Wechsler Intelligence Scale for Children-Revised (WISC-R) Digit Span Test, Card sorting Test<br>• Stanford-Binet Memory for Sentences and Digit String test<br>• Recall and recognition test |

### 2.2.4 Considerations for Study Quality Evaluation

This section summarizes how specific study characteristics factored in overall quality category. [Note: these study quality considerations are specific to issue of relevance to this document, namely potential for neurodevelopmental effects of OPs. These considerations are considered 'fit for purpose' under this context and could differ in another regulatory or scientific context.] The literature base evaluated is heterogeneous, as noted in Section 2.2.3, consisting of various study designs implemented in U.S. and foreign study populations. Pesticide exposure assessments variously relied on, for example, exposure biomarkers, maternal self-reports, and other proxy indicators of OP pesticide exposures. Outcome assessments were similarly varied, relying, for example, on biomarkers of biological effects, birth records, maternal self-reports, and clinical instruments designed to evaluate neurocognitive and neurobehavioral development. These design elements have potential impacts on study quality and relevance to this document. Each study was therefore judged to be of high, moderate, or low quality in each of the following six domains effecting study quality: Study design, exposure assessment, outcome assessment, confounder control, statistical analysis, susceptibility to bias (See Table 2.2.4-1 for general considerations under each domain).

#### 2.2.4.1 Study Designs

Four basic study designs were used in the literature reviewed for this document: cohort study, case-control study, cross-sectional study, and ecologic study. The first of these two constitute the two basic types of observational (i.e. non-interventional) studies used to evaluate relative incidence of health and disease outcomes by exposure status. The latter two are generally considered descriptive or hypothesis generating study designs, though they too can be used to test hypotheses regarding relative prevalence of health outcomes and, under certain conditions, incidence as well.

**Table 2.2.4-1 Study Quality Considerations**

| Parameter | High (Score 8-12) | Moderate (Score 4-7) | Low (Score 0-3) |
|---|---|---|---|
| Study Design | Prospective, exposure precedes disease | Case Control | Cross sectional<br><br>Ecological |
| Exposure assessment | Exposure assessment includes information on specific OP a.i.'s (e.g., CPF, MAL), or urinary metabolite (TCPy, IMPy,), or high quality questionnaire based chemical specific exposure assessment during relevant exposure | Non-specific biomarker of exposure (DAP), or effect (AChE/BuChE), or questionnaire based individual level information on the OP class, or sub-class | Low quality questionnaire based exposure assessment, or ecologic exposure assessment, with or without validation |

| | | | |
|---|---|---|---|
| | window (pre-natal, early life) | | |
| Outcome Assessment | Standardized tool, validated in study population; or, medical record review with trained staff (birth characteristics) | Standardized tool, not validated in population, or screening tool; or, medical record review, methods unstated (birth characteristics) | Selected sections of test, or maternal report, other; or, maternal/paternal self-report (birth characteristics) |
| Confounder control | Good control for important confounders relevant to OP-ND question, and standard confounders | Moderately good control confounders, standard variables, not all variables for OP-ND question | Multi-variable analysis not performed, no adjustments |
| Statistical Analysis | Appropriate to study question and design, supported by adequate sample size, maximizing use of data, reported well (not selective) | Acceptable methods, questionable study power (especially sub-analyses), analytic choices that lose information, not reported clearly | Minimal attention to statistical analyses, comparisons not performed or described clearly |
| Risk of (other) bias (selection, differential misclassification, other) | Major sources of other potential biases not likely present, present but analyzed, unlikely to influence magnitude and direction of the risk estimate | Other sources of bias present, acknowledged but not addressed in study, may influence magnitude but not direction of estimate | Major study biases present, unacknowledged or unaddressed in study, cannot exclude other explanations for study finding |

(Adapted from Munoz-Quezada *et al.* 2013)

## Cohort Study

A commonly used design in this literature was the cohort study (See Tables 2.2.5-1, 2.2.5-2, and Appendix 3 for examples). In a typical cohort study, individuals are classified according to exposure status (i.e., presence, absence, or magnitude of exposure), and then followed over time to quantify and compare the development (i.e., incidence) of the health outcome of interest by exposure group. Conceptually, the non-exposed comparison group in a cohort study provides an estimate of the incidence of the outcome among the exposed, had they, counter-to-fact, not been exposed. Apart from chance variations, a valid cohort study comparing exposed individuals to non-exposed individuals provides an estimate of the relative risk (or rate) of the disease associated with exposure. Ideally, the exposed and non-exposed groups are exchangeable, in the sense that switching the exposed to non-exposed, and non-exposed to exposed would yield the same measure of association (e.g., relative risk). If this were the case then, apart from chance, a cohort study would yield a measure of association equivalent to that

produced in a corresponding (intervention) study where exposure status was randomly assigned.

The chief advantage of the cohort study design is that it affords the investigator the opportunity to avoid and/or adjust for potential biases (i.e., selection bias, information bias, and confounding). Cohort studies also allow for discernment of the chronological relationship between exposure and outcome, and can be particularly efficient for studying uncommon exposures. The primary disadvantage of the cohort study design is logistical inefficiency with respect to the necessary time, expense, and other resources needed to conduct them. Cohort studies are particularly inefficient for evaluating associations with rare outcomes and diseases with long induction or latency periods. Though prospective studies are often logistically less efficient relative to other study designs (e.g., the case-control study), these logistical concerns can be minimized in cohort studies of short duration, such as those used to evaluate prenatal OP pesticide exposure effects on birth outcomes or other outcomes of neonatal development assessed shortly after birth.

Two sub-categories of cohort studies – prospective and retrospective - are often applied to distinguish between studies in which the health outcome has occurred (retrospective study), or has not occurred (prospective study) at the time the investigators initiate the study. This distinction is important primarily as it pertains to the potential differences in the quality (e.g., completeness, accuracy, and precision) of information that can be ascertained by the investigators, and also as it relates to potential sources of bias. Although not always true, the prospective study design is considered the preferable of the two, as investigators can potentially have more choices in determining how exposure, outcome, and covariate information is collected. In a retrospective study conducted to evaluate the same hypothesis, by contrast, the investigators would have to rely on exposure information such as maternal self-report. Such reporting is subject to (human) errors in recall. Moreover, the outcome status of the child (i.e., whether a child has developmental delays that are known to the mother) may influence the recall of prenatal OP pesticide exposure by the mother.

**Case-control Study**

In a typical case-control study (see, for example, Dawbrowski *et al*., 2003 in Appendix 3 and Shelton *et al*. 2014 in Table 2.2.5-2), individuals are classified according to their outcome status (i.e., cases who have developed the outcome of interest, and controls who represent the population from which the cases arise). The relative odds of exposure are then compared between cases and controls. The primary advantage of case-control studies is that they are logistically efficient relative to cohort studies. In fact, properly conducted case-control study can be conceptualized as a cohort study with efficient sampling of exposure among the cohort, yet they can often be conducted at a fraction of the cost, in a fraction of the time as a corresponding cohort study. Case-control studies can be used to examine associations between multiple exposures and a given health outcome. They are particularly efficient for evaluating rare outcomes but are inefficient for studying uncommon exposures. The primary weakness of the case-control study is the potential for selection bias, which arises if the exposure distribution among the control subjects is not representative of the exposure distribution

among the population that gave rise to the cases. Case-control studies that rely on self-reported exposure measures are also susceptible to information bias.

**Cross-sectional study**

Cross-sectional studies (see, for example, Suarez-Lopez *et al*. in Table 2.2.5-2 and Grandjean *et al*. 2006 in Appendix 3) are used to evaluate associations between exposure and outcome prevalence in a population at a single point in (or period of) time. The primary advantage of a cross-sectional study is logistical efficiency; they are relatively quick and inexpensive to conduct, as a long period of follow-up is not required, and exposure and outcome assessments occur simultaneously. Cross sectional studies have three primary *potential* disadvantages: 1) potential difficulty in discerning the temporal relationships (i.e., whether the exposure precedes the outcome); 2) estimating outcome prevalence rather than incidence of the outcome; and 3) the possible overrepresentation of cases of the outcome with long duration relative to the average in the population, and often with a better prognosis.

**Ecological study**

Ecological studies are used to evaluate associations between exposures and outcomes using population-level rather than individual-level data. For example, Nevison (2014, Appendix 3) uses annual estimates of pesticides applied to crops and population level autism prevalence to assess the association between OP pesticide exposure and autism. The primary advantages of ecological studies are related to logistical efficiency, as they often rely on pre-existing data sources and require no individual-level exposure, outcome, or covariate assessments. The primary weakness of the ecologic study is the potential for confounding and resultant inappropriate extrapolation of associations observed on the aggregate-level to associations on an individual level. The mistaken belief that associations observed at the population level exist at the individual level is referred to as the ecological fallacy.

In judging an individual study's contribution to the strength of evidence in the epidemiologic literature base, the following hierarchy of observational study designs was considered (from most to least preferred): prospective cohort study, retrospective cohort study, case-control study, cross-sectional study, ecological study. It is important to note, however, that this hierarchy of study designs reflects the *potential* for the collection of high quality information (related to exposure, outcome, confounders, and effect modifiers) and *potential* for efficient and valid estimation of the true association. Thus, in deliberating on quality, care has been taken to consider the circumstances and particulars of each individual study. For example, a well-conducted case-control study of a rare outcome can provide much higher quality evidence vis-à-vis the association of interest than a poorly conducted prospective cohort study of the same relationship. For this report, the placement of the study design in the aforementioned hierarchy of observational study designs was but one facet of the judgment of study design quality. Additional consideration was given to whether the study was *well conducted*, independent of study design type. The particulars of a study's design, specifically the design elements employed to minimize and adjust for biases, were also considered. Finally, the relevance of each study with respect to the association of primary interest in this initiative,

namely the relationship between prenatal (and early life) OP pesticide exposures and fetal and child neurodevelopment was considered.

#### 2.2.4.2 Exposure Measures

There were three major categories of exposure assessment employed in this literature: exposure biomarkers, participant-reported proxy exposure via questionnaire, and objectively obtained proxy indicators of exposure. Although one study included environmental wipe sampling results in the exposure assessment, urinary biomarker measures were also included and no differentiation of the two approaches was presented, so this exposure assessment category has not been included. The merits and the disadvantages of the three primary exposure assessment strategies are discussed below.

Most of the studies reviewed herein assessed biomarkers of exposure quantified in samples of biological media (most often urine, but also blood, serum, and breast milk). These biomarkers were of three types: 1) OP pesticide residues, 2) metabolites of specific OP pesticides, and 3) non-specific OP metabolites.

The most commonly measured biomarkers were urinary dialkylphosphate metabolites (DAP). These non-specific markers are easily quantified using gas chromatography/mass spectrometry and related methods. Though objective, use of urinary DAPs as biomarkers for OP pesticide exposures has limitations, including substantial temporal variability, often varying substantially over short time scales (i.e., day-to-day). Quantification of DAPs in a single urine sample may not represent an individual's usual exposure to OP pesticides over the time period of interest (e.g., pregnancy) in their utility as biomarkers of OP pesticide exposure. Urinary DAP metabolite levels may also reflect exposure to ambient metabolites in addition to exposure to OP parent compounds. In this literature, errors in DAP as a biomarker of OP pesticide exposure are likely to be non-differential with respect to outcomes. Epidemiologists often distinguish between two mechanisms or types of misclassification – those that are non-differential (or random) and those that are differential (non-random). See Section 2.2.4.6 for further discussion of these exposure misclassifications.

Many studies used questionnaire-based exposure assessments in which study participants (typically mothers) self-reported their exposures, in addition to, or instead of, quantifying OP pesticide biomarkers in samples of biological media. These exposure assessments typically include querying OP pesticide exposure directly, or asked study participants to report on behaviors and conditions associated with pesticide use (e.g., occupation, tasks). Such reporting likely misclassifies actual OP pesticide exposure. If conducted as part of a prospective exposure assessment, these errors are likely to be non-differential with respect to the outcome(s) of interest. In the context of a retrospective assessment in which the mother has knowledge of the outcome status of the child, these errors may be differential or non-differential.

Several studies used proxy measures (including ecological indicators) of pesticide exposure. These included, for example questions about occupational use and exposure to pesticides, distance from residence to fields where pesticides were applied, the proportion of land in a

specified area dedicated to agricultural uses, and occurrence of pregnancy during the pesticide application season. Again, substantial non-differential exposure measurement error/misclassification of exposure is likely.

For this evaluation, studies employing exposure assessments that quantified biomarkers of specific OP pesticides (e.g., chlorpyrifos, malathion, diazinon), or urinary metabolites of these pesticides (e.g., TCPy, malathion dicarboxylic acid, 2-isopropyl-4-methyl-6-hydroxypyrimidine), or high-quality chemical-specific exposure quantitation during relevant exposure-time windows (i.e. pre-natal, early life) were given the highest weight. Studies that quantified levels of non-specific biomarker of OP pesticide exposure (e.g., DAPs), or exposure effects (e.g., AChE, BuChE,) were given a moderate weight. Studies relying on high-quality survey-based individual-level information on pesticide exposure were also assigned a moderate weight. Exposure assessments that only crudely or subjectively classified pesticide exposures and ecologic and other proxy measures of exposure were assigned a low weight.

Two studies focusing on neurodevelopmental outcomes measured OP exposure using both DAPs and malathion dicarboxylic acid (a metabolite of malathion) (Eskenazi *et al.*, 2007; Engel *et al*., 2007), with TCPy exposure also being measured in one of these studies (Eskenazi *et al.*, 2007). Additionally, two studies focusing on birth outcomes measured OP exposure by testing for specific OPs, with Whyatt *et al.* (2004) measuring chlorpyrifos and diazinon; and Eskenazi *et al*. (2004) testing for DAPs and for seven pesticide specific metabolites (MDA - derived from malathion; PNP - derived from methyl parathion, parathion, and other nonpesticide chemicals; TCPy - from chlorpyrifos and chlorpyrifos methyl; DEAMPY - from pirimiphos methyl; IMPY - from diazinon; CMHC - from coumaphos and coumaphos methyl; CIT - from isazophos and isazophos methyl). Finally, several method validation studies tested for specific OP pesticides, but did not evaluate the association between these exposures and specific adverse health outcomes (Whyatt *et al.*, 2007; Whyatt *et al.*, 2009; Bradman *et al.*, 2003).

With the exception of the studies discussed above and those focusing exclusively on TCPy or chlorpyrifos, the majority of the epidemiological studies focus on the association of OPs exposure and various neurodevelopmental outcomes. The OP exposure being assessed in these studies used concentrations of urinary dialkyl phosphate metabolites (DAPs). Total DAPs is a non-specific measure of OP exposure and is the sum of six separate molecules - three dimethyl alkylphosphate (DMAP) molecules of DMP, DMTP, DMDTP, and three diethyl alkylphosphate (DEAP) molecules of DEP, DETP, and DEDTP. Each metabolite is a breakdown product from multiple OPs (Table 2.2.4-2). Specifically, DMP, DMTP, and DMDTP are associated with 18, 13, and 5 OPs, whereas DEP, DETP, and DEDTP are associated with 10, 10, and 4 OPs, respectively. Thus, using DAPs as an exposure measure, it is not possible to separate the exposure and associated effects for single, specific OPs. For studies evaluating TCPy (e.g., Fortenberry *et al*., 2014; Eskenazi *et al.*, 2007; Whyatt *et al.*, 2009), this molecule is a metabolite of chlorpyrifos, chlorpyrifos-methyl, and the herbicide triclopyr. TCPy is a primary environmental degradate of chlorpyrifos, chlorpyrifos-methyl, and triclopyr, and is found on food treated with these pesticides. Studies focusing solely on chlorpyrifos could assess exposure to only this OP (e.g., Lovasi *et al*., 2010; Whyatt *et al*., 2004; Rauh *et al*., 2011).

**Table 2.2.4-2 CDC Table of Organophosphate Pesticides and Their Dialkyl Phosphate Metabolites (2008)**

| Pesticide | DMP | DMTP | DMDTP | DEP | DETP | DEDTP |
|---|---|---|---|---|---|---|
| Azinphos methyl | \| | \| | \| | | | |
| Chlorethoxyphos | | | | \| | \| | |
| Chlorpyrifos | | | | \| | \| | |
| Chlorpyrifos methyl | \| | \| | | | | |
| Coumaphos | | | | \| | \| | |
| Dichlorvos (DDVP) | \| | | | | | |
| Diazinon | | | | \| | \| | |
| Dicrotophos | \| | | | | | |
| Dimethoate | \| | \| | \| | | | |
| Disulfoton | | | | \| | \| | \| |
| Ethion | | | | \| | \| | \| |
| Fenitrothion | \| | \| | | | | |
| Fenthion | \| | \| | | | | |
| Isazaphos-methyl | \| | \| | | | | |
| Malathion | \| | \| | \| | | | |
| Methidathion | \| | \| | \| | | | |
| Methyl parathion | \| | \| | | | | |
| Naled | \| | | | | | |
| Oxydemeton-methyl | \| | \| | | | | |
| Parathion | | | | \| | \| | |
| Phorate | | | | \| | \| | \| |
| Phosmet | \| | \| | \| | | | |
| Pirimiphos-methyl | \| | \| | | | | |
| Sulfotepp | | | | \| | \| | |
| Temephos | \| | \| | | | | |
| Terbufos | | | | \| | \| | \| |
| Tetrachlorviphos | \| | | | | | |
| Trichlorfon | \| | | | | | |

DMP = dimethylphosphate; DEP = diethylphosphate; DMTP = dimethylthiophosphate; DMDTP = dimethyldithiophosphate; DETP = diethylthiophosphate; DEDTP = diethyldithiophosphate.

The CHARGE study (Shelton *et al.*, 2015) used a different method for exposure assessment. This study used geospatial analysis to focus on the residential proximity to OP exposure and the association of this exposure with autism spectrum disorders. OP exposure was assessed by Shelton *et al* (2015) using data from the California Department of Pesticide Regulation, with five OPs accounting for a total of 73% of the exposure and each accounting for 10% or more of the exposure (chlorpyrifos, acephate, diazinon, bensulide, and dimethoate); eight OPs

accounting for a total of 25% of the exposure and each accounting for 1% or more of the exposure (malathion, methyl parathion, azinphos-methyl, phosmet, oxydemeton-methyl, ethephon, naled, and methidathion); and eight OPs accounting for a total of 2% of the exposure and each accounting for 0.1% or more of the exposure (methamidophos, phorate, disulfoton, fenamiphos, coumaphos, parathion, ethoprop, and sulfotep).

#### 2.2.4.3 Neurological and Other Outcome Measures

With some exceptions, the outcomes assessed in this literature fall into three broad categories: 1) neurobehavioral and/or neurodevelopmental status; 2) birth outcomes; and 3) neurodevelopmental diseases and/or disorders (see Table 2.2-b).

There is a broad body of literature available on the use and interpretation of instruments designed to quantify neurological status that is beyond the scope of this summary. Many instruments were used in this literature to assess infant neurodevelopment and neurobehavior (see Table 2.2-b). Importantly, performance on some of these tests can be influenced by the administrator of the assessment. Also of concern is whether these assessments are sensitive enough to distinguish potentially subtle effects of OP pesticide exposure. Some degree of error in the assessment of neurological outcomes is likely in all of the studies, though the errors were unlikely to be related to exposure status in the well conducted studies.

The assessment of birth outcomes in this literature was primarily conducted by reviewing of medical records or birth certificates; these assessments are likely to have minimal errors, and errors that do arise are almost certainly non-differential with respect to exposure status. In some studies, however, birth outcomes were reported by the mother and in these instances, differential misclassification is possible.

The studies that evaluated specific diagnoses, autism spectrum disorders or developmental delay, for example, relied on existing medical records, with some effort to validate diagnoses in a subset of the investigations.

Studies that relied on standardized, validated instruments to assess neurodevelopment, medical records of birth outcomes, or validated diagnosis of disease states were weighted highly in the judgment of study quality. Those that used standardized instruments that had not been validated in the relevant population or screening assessments were given a moderate weight, while studies relying on selected sections of neurodevelopment assessment tools, maternal report of outcome status, or aggregated (ecological) outcome measures were given the lowest weight.

#### 2.2.4.4 Statistical Analysis

Statistical analyses that were appropriate to the study question and study design, supported by adequate sample size, maximized the use of available data, and were well characterized in the report were weighted most highly. Acceptable statistical methods, moderate study power, and analytic choices that resulted in the loss of information or that were not clearly reported were

given moderate weight. Reports with only minimal attention paid to the conduct and reporting of the statistical analyses were given the lowest weight.

#### 2.2.4.5 Confounding

Risk factors for early life neurodevelopment perturbations that are associated with OP pesticide exposure, but not caused by pesticide exposure, are potential confounders in this literature. Socioeconomic determinants of child development (including, for example, maternal education and access to prenatal and early life health care), quantity and quality of parent-child interactions, and exposure to other environmental toxicants (e.g., lead, PCBs, other pesticides) are difficult to measure and were either not accounted for, or inadequately accounted for in many studies. That said, some studies were relatively homogeneous with respect to these factors and thus limited confounding by design, while others attempted to quantify these factors and adjust for them. Other important potential confounders, such as the child's sex, are easy to identify and adjust for analytically.

#### 2.2.4.6 Risk of Bias

The internal validity of the studies reviewed was judged by noting the design strategies and analytic methods used in each study to constrain or eliminate selection bias, information bias, and confounding. Selection bias occurs when the sampling of the population by the investigator yields a study population that is not representative of the exposure and outcome distributions in the population sampled. Put simply, selection bias occurs if selection of the study sample yields a different estimate of the measure of association than that which would have been obtained, had the entire target population been evaluated. Although there are numerous sources of selection bias, there are several mechanisms that may have induced selection bias in the studies reviewed: less than 100% participation rates of eligible individuals due to non-responsiveness or refusal (self-selection bias); loss to follow-up (i.e. failure to retain all study participants initially enrolled in the study); and, in a case-control study, control selection bias arising because the exposure distribution in the control sample does not represent the exposure distribution of the study base (i.e., the population that gave rise to the cases or more formally, the person-time experience of that population).

Information bias (also referred to as observation bias) arises when study participants are incorrectly categorized with respect to their exposure or outcome status, or when errors arise in the measurement of exposure or outcome, in the case of continuously distributed measures. Epidemiologists often distinguish between two mechanisms or types of misclassification – those that are non-differential (or random) and those that are differential (non-random). Non-differential misclassification of exposure (or non-differential exposure measurement error) occurs when the probability or magnitude of error in the classification or measurement of exposure is independent of the outcome status of the study participants. Similarly, non-differential misclassification of outcome (or outcome measurement error) occurs when the probability or magnitude of error in the assignment of outcome status or level is independent of exposure status. In contrast, differential exposure misclassification (or measurement error) occurs when the error in the exposure assignment is not independent of the outcome status.

The mechanisms that cause non-differential misclassification in this literature include errors in the medical records, laboratory errors, sampling of biospecimens for biomarker assays, and errors in recall. The mechanisms that induce differential misclassification include recall bias, and interviewer/observer bias. Note that mismeasurement of confounders can result in residual confounding of the association of interest, even when adjustment for that confounder has been conducted in the analysis.

Studies in which major sources of potential biases were not likely to be present, or in which potential sources of bias were present but effectively addressed and analyzed to maximize the study validity, and those in which sources of bias were unlikely to influence the magnitude and direction of the risk estimate were given a high weight. Studies where sources of bias were present and acknowledged by the authors but not addressed in the study and yet may influence the magnitude, but not direction of the association estimate received a moderate weight. A low weight was given to studies in which major biases were present and yet were not acknowledged or addressed in the study, such that they cannot be excluded as an alternative explanation for the study finding.

### 2.2.5 Review of Quality Results

Each of the studies in the 2015 review was judged to be of high, moderate, or low quality in each of five domains of study design and methodology effecting study quality as discussed above in Section 2.1. The results of the quality assessment are presented separately below for each group. The quality categories represent to the total evaluation. In Section 2.2.6, further evaluation of the study design and exposure assessment of the medium and high quality studies. This further evaluation led to additional studies being removed from the final analysis, with these excluded studies not being considered further in the remaining sections of this document.

#### 2.2.5.1 "High" Quality Group

Six studies (8 articles) were assigned a high quality rating, as shown in Table 2.2.5-1. In general these were prospective birth cohort studies with moderate to high sample size; exposure assessment was based on an objective biomarker measure, the outcome measurement(s) utilized standardized tests and trained data collectors, appropriate statistical analyses were performed, considering relevant covariates, and risks of bias were minimized to the extent possible. For example, Fortenberry *et al*. (2014, Table 2.2.5-1) reported on findings in the ELEMENT study population of 187 mother-child pairs, assessed third trimester maternal urinary TCPy as a biomarker of prenatal chlorpyrifos exposure, and employed a trained and experienced research team to administer a battery of validated ADHD/psychometric assessments (Conner's' Parental Rating Scales-Revised, Conner's' Continuous Performance Test, and the Behavior Assessment System for Children-Parental Rating Scales) that had been translated into Spanish.

**Table 2.2.5-1. High Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Barr *et al*. (2010) | New Brunswick, New Jersey | Mother who underwent elective Cesarean delivery at term and their newborns, Saint Peter's University Hospital | Cross sectional birth cohort, modest sample numbers<br><br>N=150<br><br>Mother-infant pairs | Objective measure, prenatal - parent CPF in maternal and cord serum at birth - exposures do not necessarily precede outcome | Birth outcomes from medical records | Generally appropriate; missing a few, e.g., maternal age, SES, nutrition, pre-natal care, race | Appropriate multivariate analysis | Convenience sample – deemed low probability of selection bias; low potential for exposure or outcome misclassification |
| Bouchard *et al*. (2010) | U.S. National Population | National Health and Nutrition Examination Survey (NHANES) | Cross-sectional, large sample size<br><br>N=1139<br><br>Age 8-15 years | Objective measure, postnatal – single non-OP specific child urinary DAPs | ADHD from DISC-IV diagnosis or medication in children 8-15 years old, standard protocol, trained interviewers | Appropriate: also included blood lead | Appropriate, accounted for NHANES multistage probability sampling | Low probability of selection bias; some potential for non-differential misclassification of exposure |
| Fortenberry *et al*. (2014) | Mexico City, Mexico | Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) study (3 cohorts 1994-1997 (cohort 1), 1997-2000 (cohort 2), and 2001-2005 (cohort 3)) | Prospective Birth Cohort Study – modest sample size<br><br>N=187<br><br>Mother-infant pairs | Objective measure - TCPy concentration was measured in third trimester maternal urine samples. In a subset of women randomly selected urinary TCPy in samples collected during all three trimesters of pregnancy was measured. | ADHD - LP - Conner's' Parent Rating Scales-Revised (CPRS-R), Conner's' Continuous Performance Test (CPT), and Behavior Assessment System for Children-Parental Rating Scales (BASC-PRS) – These are screening tools, not diagnostic tools. Standardization and quality control checks were conducted by reviews of videotaped evaluations. | Appropriate. Included continuous maternal IQ, education, socioeconomic status and blood lead one month after delivery, breast feeding (yes/no), child's sex, continuous age at testing, birth length and head circumference at birth. | Appropriate. Multivariable linear regression | History of maternal and paternal exposure to pesticides not included, use of a single urinary measure to estimate exposure, potential for differential exposure misclassification possible as mothers were likely aware of the neurobehavioral status of their children |
| Fortenberry *et al*. (2014a) | Mexico City, Mexico | Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) study | Prospective Birth Cohort Study – large sample size<br><br>N=591<br><br>Mother-infant pairs | Objective measure - blood samples from mother–child pairs were analyzed for PON1. OP exposure was not explicitly assessed. | ADHD - LP - Conner's' Parent Rating Scales-Revised (CPRS-R), Conner's' Continuous Performance Test (CPT), and Behavior Assessment System for Children-Parental Rating Scales (BASC-PRS) - These are screening tools, not diagnostic tools. | Appropriate: Included continuous maternal IQ, education, socioeconomic status and blood lead one month after delivery, breast feeding (yes/no), child's sex, continuous age at testing, birth length and head circumference at birth. | Appropriate. Multivariable linear regression | Potential for differential exposure misclassification possible as mothers were likely aware of the neurobehavioral status of their children. |

**Table 2.2.5-1. High Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Contro | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Quirós-Alcalá *et al*. (2011) | Salinas Valley, California, USA | Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) | Prospective Birth Cohort Study Outcomes assessed at:<br>N=149 (6 months)<br>N=149 (1 year)<br>N=97 (3.5 years)<br>N= 274 (5 years) | Objective biomarker, prenatal OP pesticide exposure (DAP) quantified in two maternal spot urine samples at 14 and 27 weeks gestation (on average), and in single child urine sample provided at time of each outcome assessment | Autonomic nervous system dysregulation at ages 6 months, 1 year, 3.5 years, and 5 years; standard protocol, trained, bilingual research staff, appropriate to age of child. May not be sensitive to resolve subtle effects of OP pesticide exposure. | Appropriate. Included maternal and child demographic and SES indicators (age, education). | Appropriate multivariate analysis | Selection bias possible due to considerable loss to follow-up; Residual confounding likely small in magnitude; some potential for non-differential misclassification of exposure (possible explanation for null findings). |
| Rauch *et al*. (2012) | Cincinnati, Ohio, USA | Health Outcomes and Measure of the Environment (HOME) | Prospective Birth Cohort Study – large sample size.<br><br>N=306<br><br>Age: newborns | Objective biomarker of prenatal OP pesticide exposure (DAP) quantified in two maternal spot urine samples provided at 16 and 26 weeks gestation | Abstracted birth weight from medical records, calculated gestational age from mother's self-reported date of last menstrual period. When gestational age not available, results of an ultrasound was used or a Ballard examination performed just after delivery. | Appropriate. Included income, education, maternal depressive symptoms, maternal IQ, insurance status, area of residence, prenatal care, PON1 genotype, and gestational exposure to alcohol, lead, and tobacco. | Appropriate. multivariable regression | Unable to rule out the possibility that differences in DAP concentrations partially reflect individual variation in metabolism, Recall error |
| Wolff *et al*. (2007) | New York City, U.S. | Mount Sinai Children's Environmental Health Study | Prospective Cohort – moderate sample size<br><br>N=404<br><br>Mother-infant pairs | Objective biomarker, prenatal OP pesticide exposure (DAP), malathion (MDA) quantified in single maternal spot urine sample provided during the 3rd trimester | Birth outcomes from computerized perinatal database | Appropriate: different covariates for each outcome, e.g., weight: maternal age, race/ethnicity, maternal BMI*pregnancy weight gain, infant sex, gestational age, creatinine | Appropriate PROC GLM, with varying covariates | Selection bias unlikely; some potential for non-differential misclassification of exposure |
| Furlong *et al*. (2014) | New York City,<br><br>U.S. | Mount Sinai Children's Environmental Health Study | Prospective Cohort<br><br>N=136<br><br>Age 7-9 years | Objective biomarker, prenatal OP pesticide exposure (DAP) quantified in single maternal spot urine sample provided during the 3rd trimester, simple imputation < lowest level of detection (LLOD) | Reciprocal social impairment at age 7-9 years assessed using the Social Responsiveness Scale, completed by mothers. Designed to assess reciprocal social behaviors in evaluating ASDs, used here as general indicator of impaired social responsiveness. | Appropriate. Included maternal and child demographic and SES indicators (age, education) and ETS (not other environmental toxicants). | Appropriate multivariate analysis | Selection bias possible due to considerable loss to follow-up; Residual confounding likely small in magnitude; some potential for non-differential misclassification of exposure possible explanation for null findings). |

### 2.2.5.2 "Moderate" Quality Group

Eleven studies (15 articles) were assigned a moderate quality rating, as shown in Table 2.2.5-2. In general, these were cross-sectional or prospective cohort studies with small to high sample size; exposure assessment was based on a non-specific biomarker measure or current self-report, the outcome measurement(s) utilized standardized tests or screening tools, appropriate statistical analyses were performed, considering some but maybe not all relevant covariates, and risks of bias were minimized to some extent. For example, Guodong *et al*. (2012) cross-sectionally evaluated the relationship between DAPs concentrations in urine sampled from 301 young children as an objective, non-specific marker of prenatal OP pesticide exposure and Developmental Quotients based on the Gesell Developmental Schedules adapted for a Chinese population.

**Table 2.2.5-2. Moderate Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Andersen *et al.*(2015) | Denmark | School age children living in Denmark | Prospective Birth Cohort, small sample size<br><br>N=177<br><br>(112 maternal occupational pesticide exposure during pregnancy; 65 without)<br><br>Age: 6-11 years | Prenatal occupational pesticide exposure ascertained by maternal interview at enrollment | Objective standardized clinical exam; neurophysiological status (heart rate variability); Previously validated, neuropsychological testing with demonstrated sensitivity to environmental pollutants; administered and scored by single neuropsychologist | Appropriate. Self-reported by mother. Included age, maternal demographics and risk factors, SES indicator (broad categories), maternal smoking and alcohol use. Possibility of recall errors (residual confounding). Self-reported (by mother) | Appropriate multivariate analysis. Evaluated numerous hypotheses without adjustment for multiple comparisons. Also constructed structural equation model of interaction between child sex and prenatal pesticide exposure on general intellectual ability (parameterized as a latent variable) | Selection bias possible due to loss to follow-up; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |
| Guodong *et al*. (2012) | Shanghai, China | 2-year old children visiting community hospitals | Cross-sectional, large sample size<br><br>N=301<br><br>Age 23-25 months | Objective biomarker, OP exposure (DAP) quantified in single child spot urine sample, simple imputation <LLOD; OP exposure also assessed via questionnaire administered to mothers after delivery | Developmental Quotients based on Gesell Developmental Schedules, adapted for Chinese population | Appropriate. Assessed via mother report via questionnaire, included child sex, maternal demographics and risk factors, SES indicators (maternal education, household income), maternal smoking and alcohol use; Possibility of recall errors | Appropriate multivariate analysis. Assumed linear relationship between log-transformed DAP level and DQ Scores | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |
| Handal *et al*. (2007) | Cayambe-Tabacundo region, Ecuador | EcoSalud Project: infants and young children 3 to 61 months of age in lower-altitude communities A&B dominated largely by cut-flower production and in more traditional rural, higher altitude community C. | Cross-sectional, modest sample size from census<br><br>N=283<br><br>Age 3-61 months<br><br>- 154 high exposure<br><br>- 129 low exposure | Objective proxy measure: Community of residence | Ages and Stages Questionnaire (ASQ), adapted into local vernacular, 2 trained testers – considered a screening tool | Appropriate: child health status (anemia, stunting) and other characteristics of the home environment (stimulation by 2 methods) | Appropriate: Multiple linear regressions to evaluate associations between community of residence and delayed development, Pairwise t-tests and chi-square to assess mean difference in ASQ score; Effect size Cohen's d | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |
| Handal *et al*. (2008) | Cayambe-Tabacundo region, Ecuador | EcoSalud Project: Children 3 to 23 months of age in lower-altitude communities A&B dominated largely by cut-flower production and in more traditional rural, higher altitude community C. | Cross-sectional, modest sample size<br><br>N:121<br><br>Age 3-23 months | Proxy: Distance home to farm, parental employment, pesticide use on domestic crops & within home, child play activities | Ages and Stages Questionnaire (ASQ) (ages 24-61 months), a screening tool; Visual Motor Integration (VMI) Test (ages 48-61 months); two trained testers. | Appropriate: As above | Appropriate: As above | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |

**Table 2.2.5-2. Moderate Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Handal *et al.*(2007B) | Cayambe-Tabacundo region, Ecuador | EcoSalud Project: Children 24 to 61 months of age in lower-altitude communities designated A & B were dominated largely by cut-flower production than in more traditional rural, higher altitude community C. | Cross-sectional, modest sample size<br><br>N:<br><br>ASQ - 142<br><br>Age 24-61 months<br><br>VMI - 57<br><br>Age 48-61 months | Proxy: Maternal employment during pregnancy, child plays outdoors, Pesticide use on domestic crops , inside home | Ages and Stages Questionnaire (ASQ) (3-23 months) – screening tool; Reach-and-grasp, UC Berkeley Preferential Looking Test Cards; trained tester | Appropriate: As above | Appropriate: As above;<br><br>Prehension- Logistic regression models, generalized estimation equations (GEE) to account for between-test dependence | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |
| Llop *et al.* (2013) | Spain | INMA (Environment and Childhood) Project | Prospective Birth Cohort, large sample size<br><br>N=1980<br>Age 14 months | Maternal self-report of prenatal and postnatal indoor pesticide use (pesticide spray or use of a plug-in device assessed via questionnaire) | Mental and psychomotor development at 14 months assessed using validated instrument (Bayley Scales of Infant Development) | Appropriate. Self-reported by mother. Included maternal demographics and risk factors (Age, BMI) SES indicators (maternal education, occupation), maternal smoking and alcohol use. Also childcare behaviors (breast feeding, number of siblings, day care) Possibility of recall errors (residual confounding) | Appropriate multivariate analysis | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure |
| Oulhote and Bouchard (2013) | Canadian National Population | Canadian Health Measures Survey (CHMS; cycle 1, 2007–2009); | Cross-sectional Study, large national sample size<br><br>n = 779 children<br><br>Age 6-11 years | Child urinary DAP, DMP, DEP collected within 2 weeks of survey questionnaire completion by the parents. | Behavioral problems in children based on the parent version of the Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997) - SDQ is a validated screening questionnaire and accepted by parents. | Appropriate: Included sex, age, race/ethnicity, family income, parental education, blood lead levels, maternal smoking during pregnancy, birth weight, maternal age at child's birth, child's BMI, and fasting status (fasting duration at urine collection > 10 hour/≤ 10 hour) | Appropriate: Logistic Regression | Selection bias unlikely; Study population not mixed, single urine sample |
| Petit *et al.* (2010) | Brittany, France | Newborn children in the PELAGIE Study. | Prospective Birth Cohort, large sample size<br><br>N= 3,159<br><br>Age: newborns | Ecological, proxy indicator of exposure (proportion of municipality devoted to agricultural activity | Objectively measured birth outcomes assessed using hospital records | Appropriate. Maternal report via questionnaire. Included maternal demographics (age, BMI) and pregnancy risk factors (gestational age, hypertension, diabetes, season of pregnancy) SES indicators (district of residence, maternal education, occupation), maternal smoking and alcohol use. Also childcare behaviors (breast feeding, number of siblings, day care); Possibility of recall errors | Appropriate multivariate analysis. Multiple comparisons were conducted for this evaluation, and the authors did not adjust for multiple testing | Selection bias unlikely; Residual confounding likely, small in magnitude; potential for substantial non-differential misclassification of exposure; Misclassification of outcomes unlikely |

**Table 2.2.5-2. Moderate Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Ruckart *et al*. (2004) | Mississippi, USA (29 counties), Ohio (one multi-family facility) | Children who were 6 years or younger in MS and OH when homes sprayed with MP | Retrospective cohort, moderate sample size in 2 states, participating/ nonparticipating similar in sex and age<br>MS: N=365 (147 exposed, 218 unexposed)<br>OH: N=287 (104 exposed, 183 unexposed)<br>Age: 1.9-12.5 years old at testing | Specific OP measures: Household methyl parathion (wipe) by approved labs, OR highest urinary PNP level in household, analyzed by CDC | Variety of neuro measures: PENTB, Parenting Stress Index (PSI), Personality Inventory for Children (PIC), Vineland Adaptive Behavior Scales (VABS) | Appropriate: Income, race, site term, ethnicity, mother's use of chemicals at work, mother health/pregnancy conditions, report that child had lead or mercury poisoning (MS only); Raw scores were child age-adjusted | Appropriate: Linear regression (continuous scores), logistic regression (dichotomous scores) | Selection bias possible due to loss of follow-up; Residual confounding likely; substantial potential for differential exposure misclassification (time between spraying and testing, frequency and duration unknown) |
| Shelton *et al*. (2014) | California, USA | CHARGE Study (3 year-old children) | Case-control, large sample size<br><br>N = 970<br><br>Age 3 years<br><br>N=486 (ASD)<br><br>N=168 (DD)<br><br>N=316 (Controls) | Proxy indicator of prenatal OP pesticide exposure (residential proximity to agricultural pesticide applications defined using ecological pesticide use data) | Clinical outcomes – validated within the study | Appropriate adjustment for demographics (place of birth, race), SES indicators (paternal education) and vitamin intake during pregnancy | Appropriate multivariate analysis. | Selection bias probable; Residual confounding likely; substantial potential for differential misclassification of exposure. Outcome misclassification unlikely. |
| Suarez-Lopez *et al*. (2012) | Pedro Moncayo County, Pichincha, Ecuador | Secondary Exposure to Pesticides Among Infants, Children and Adolescents (ESPINA) Study, plus new volunteers | Cross-sectional, moderate sample size, from prior census<br>N=277<br>Exposed: (n=158)<br>Unexposed: (n=119)<br>Age 4-9 years | Proxy: Cohabitation with flower worker >1yr, Adult questionnaire | Child AChE level using commercial kit | Appropriate: Sex, age, height-for-age, hemoglobin concentration, income, pesticide use within household lot, pesticide use by neighbors, examination date, residence, distance to nearest flower plantation | Appropriate: Multiple linear regression (continuous) and logistic regression (polychotomous variables), adjusted models | Selection bias possible due to loss but deemed low; Residual confounding likely small in magnitude; potential for differential exposure misclassification |
| Suarez-Lopez *et al*. (2013a) | Pedro Moncayo County, Pichincha, Ecuador | Secondary Exposure to Pesticides Among Infants, Children and Adolescents (ESPINA) Study, plus new volunteers | Cross-sectional, moderate sample size, from prior census<br>N=307<br>Age 4-9 years | Non-specific measures:<br>Child AChE level using commercial kit;<br>Proxy: Cohabitation with flower worker >1yr, Adult questionnaire | Heart rate (HR), Systolic blood pressure (SBP), Diastolic blood pressure (DBP) by standard methods | Appropriate: Sex, age, height-for-age, hemoglobin concentration, income, pesticide use within household lot, pesticide use by neighbors, examination date, residence, distance to nearest flower plantation | Appropriate: Multiple linear regression (continuous), adjusted. | Selection bias possible due to loss but deemed low; Residual confounding likely small in magnitude; potential for differential exposure misclassification |
| Suarez-Lopez *et al*. (2013) | Pedro Moncayo County, Pichincha, Ecuador | Secondary Exposure to Pesticides Among Infants, Children and Adolescents (ESPINA) Study, plus new volunteers | Cross-sectional, moderate sample size, from prior census<br>N=271<br>Age 4-9 years | Non-specific measures:<br>Child AChE level using commercial kit | NEPSY-II test (trained examiners): general assessment battery | Model defined a priori: hemoglobin, age, sex, race, height-for-age z score, household income, maternal education, and flower worker cohabitation status | Logistic models (dichotomous and polychotomous) and linear regression models, adjusted; Effect modification according to sex among significant associations; Some imputed values | Selection bias possible due to loss but deemed low; Residual confounding likely small in magnitude; potential for differential exposure misclassification |
| Wang *et al*. (2012) | Shanghai, China | Pregnant women and newborn children | Cross-sectional Birth Cohort Study<br>N=187<br>Age: newborns | Objective biomarker of prenatal OP pesticide exposure (DAP) in single maternal spot urine sample provided at onset of labor. Simple imputation of observations < LLOD. OP pesticide exposure also assessed via questionnaire administered after delivery | Gestational age and pre-term delivery appropriately defined and assessed using medical records | Appropriate. Included maternal anthropomorphic, demographic, and SES indicators (income, occupation), predictors of high-risk pregnancy (pregnancy weight gain, gestational age) | Appropriate multivariate analysis | Selection bias unlikely; Residual confounding likely small in magnitude; potential for differential misclassification of exposure |
| Zhang *et al*. (2014) | Shenyang, China | Newborn children in a birth cohort study | Prospective Cohort, n=249 mother-infant pairs<br>Age: newborns | Biomarker of prenatal OP pesticide exposure (DAP) quantified in single maternal spot urine sample provided at delivery. Many observations < LLOD | Neonatal neurodevelopment assessed using validated instrument (Neonatal Behavioral Neurological Assessment) by trained examiners | Appropriate. Included maternal demographic and SES indicators (age, education), predictors of high-risk pregnancy (BMI, gestational age) and environmental toxicant exposure (cord blood lead); Results stratified by sex | Appropriate multivariate analysis | Selection bias unlikely; Residual confounding likely small in magnitude; considerable potential for non-differential misclassification of exposure (unlikely to account for non-null findings) |

#### 2.2.5.3 "Low" Quality Group

Fourteen studies (15 articles) were assigned a low quality rating, as shown in Appendix 3. In general these were small or pilot studies; exposure assessment was based on a proxy measure(s), outcome measurement(s) utilized screening tools or self-report, limited statistical analyses were performed, relevant covariates were not included or discussed, or risks of bias were possible. For example, Acosta-Maldonado *et al.* (2009) conducted a cross-sectional pilot study among 54 women, only nine of whom were considered to have had prenatal exposure to pesticides (defined based on either an exposure history profile or AChE level in blood sampled at the time of admission into the hospital for delivery. The outcome assessed in this study was a standardized but partially subjective assessment of placental maturity (the Placental Maturity Index, PMI). Covariates adjustment of estimated associations between prenatal exposure and PMI in this study was minimal and limited to placental characteristics.

***The remaining sections of this document do not discuss further studies identified in the 'low' category.*** Due to limitations in these studies, they do not provide reliable information evaluating associations between OP exposure and neurodevelopmental outcomes.

### 2.2.6 Assessment of Epidemiological Studies for Relevance to Analysis

Using the criteria summarized in Section 2.2.4, a total of 38 literature articles were identified in the 2015 literature review and were judged as high, moderate, or low quality. Overall, 8 articles, 15 articles, and 15 articles were judged to be of high, moderate, or low quality, respectively. For the 23 high and moderate quality studies, additional evaluation was conducted as described in this section.

While all of the moderate quality studies had strengths including sample design and outcome assessment, six of these moderate quality studies did not have sufficient exposure assessment methods to determine whether exposure to OPs actually occurred. These studies were conducted on study populations in Spain (Llop *et al.*, 2013), Ecuador (Handal *et al.*, 2007; 2007b; 2008), Denmark (Andersen *et al.*, 2015), and France (Petit *et al.*, 2010). In these studies, participants were considered exposed or unexposed to pesticides based on non-specific exposure measures, such as self-reported occupational exposure, home pesticide spraying, and proportion of municipality devoted to agricultural activity. For all of these proxy exposure assessments, the pesticides used may have included not only OPs, but also pyrethroids, fungicides, and growth regulators. Given the uncertainty about whether OP exposure actually occurred in these studies and whether observed outcomes are associated with OP exposure or with other pesticides, these studies are excluded from further analysis.

The focus of this epidemiological literature review is on the neurodevelopmental outcomes from exposures to low levels of OPs (i.e. below exposures which would result in 10% or more AChE inhibition). Three studies conducted in Ecuador focused on child AChE inhibition and the potential association of AChE inhibition with other measures including parental occupation (Suarez-Lopez *et al.*, 2012), as well as clinical autonomic nervous system (ANS) outcomes such as blood pressure and heart rate (Suarez-Lopez *et al.*, 2013a), and neurodevelopmental outcomes (Suarez-Lopez *et al.*, 2013). The range of AChE activity levels are lower in the first

tertile (range of 1.44 to 2.93 U/mL) compared to the third tertile (range from 3.33 to 4.69 U/mL). Therefore, due to the outcomes assessed and the potentially toxic cholinergic effects that were associated with these outcomes, these studies are not considered to be relevant to this review and are not discussed further here.

In a retrospective cohort study of children exposed to methyl parathion (MP) before age 6 years in Mississippi and Ohio (Ruckart *et al*. 2004), as assessed by household urinary PNP or wipe samples, exposed children performed worse than unexposed children on a few of numerous neurobehavioral development tests conducted. Specifically, participants classified as having had MP exposure had more difficulty with short term memory and attention tasks, and parents reported more behavioral and motor skill problems, relative to unexposed children. However, upon closer inspection of the results across the MS and OH study sites, these neurobehavioral outcomes are not seen consistently. These inconsistencies may be due to differences in how the exposure occurred across the sites, including the fact that MS participants and OH participants were tested 2.5 and 4.5 years after MP spraying in the home. When comparing exposed and unexposed children using general intelligence testing, integrated visual and motor skills testing, and multistep processing, they did not see any differences. The exposure scenario associated with these observations is a critical element in assessing the utility and reliability of this study. Samples were collected in locations from OH and MS where illegal spraying of methyl parathion is known to have occurred during the 1994-1996 time period. Based on the "Revised Organophosphorous Pesticide Cumulative Risk Assessment" (USEPA, 2006), methyl parathion is known to be among the more potent OPs. It is unknown whether study participants were exposed to MP levels that would have induced cholinergic effects. Therefore, given the uncertainty around this illegal use combined with the high potency for cholinergic toxicity, the agency is not emphasizing this study further in the analysis.

In addition, one high quality study on the CHAMACOS birth cohort (Quirós-Alcalá *et al*. 2011) assessed the association between DAPs and autonomic nervous system (ANS) outcomes at ages 6 months, 1 year, 3.5 years, and 5 years. The ANS outcomes assessed included heart rate and respiratory sinus arrhythmia. Overall, while there was some evidence of ANS dysregulation for infants at 6 months, these results were not consistently observed for the other assessed child (1 year, 3.5 years, and 5 years) and maternal OP exposures. There is not a body of literature to compare these results against, making it difficult to put them into context. Furthermore, this study did not focus on neurodevelopmental outcomes, which is the focus of this analysis. Consequently, this Quirós-Alcalá *et al*. 2011 study is not being evaluated further.

### 2.2.7 Birth Outcome Epidemiologic Studies

Four identified studies, three high quality and one moderate quality, focused solely on OP exposure and adverse birth outcomes related to fetal growth (Barr *et al*., 2010; Rauch *et al*., 2012; Wang *et al*., 2012; Wolff *et al.*, 2007). The birth outcomes assessed included birth weight, birth length, head circumference at birth, and gestational age. The exposure assessment for these studies was conducted using objective measures or biomarkers such as maternal urinary DAPs (Rauch *et al*., 2012; Wang *et al*., 2012; Wolff *et al*., 2007), maternal

urinary malathion (MDA) (Wolff *et al.*, 2007), or maternal and cord blood levels for specific OPs (Barr *et al*., 2010). The results from these studies are generally inconsistent, with some studies documenting statistically significant associations between OP exposure and birth outcomes whereas others do not. For example, maternal and cord blood serum levels of TCPy were inversely associated in univariate analyses conducted among participants in a prospective cohort study conducted in New Jersey (Barr *et al*. 2010), though the associations did not persist after adjusting for gravidity, maternal pre-pregnancy BMI, infant sex, and gestational age. Chlorpyrifos levels were near the lower limit of detection in this study, though detectable in 98.6% of maternal serum and 62.8% of cord serum samples. Overall, birth length was not associated with third-trimester DAP in the Mount Sinai Cohort (Wolff *et al*. 2007). However, among those with slow-activity paraoxonase-1 (PON1) or PON192, urinary total DMP (but not total DAP or DEP) was statistically significantly associated with shorter birth length (p= 0.032). Birth length was also not associated with maternal urinary DAP sampled at delivery in the cross-sectional investigation conducted in Shanghai, China (Wang *et al*. 2012). Similarly inconsistent results were observed across these studies for birth weight, gestational age, and head circumference at birth.

Overall, in this 2015 literature review, inconsistent evidence of OP exposure and association with adverse birth outcomes/fetal growth was observed (Barr *et al*., 2010; Rauch *et al.*, 2012; Wang *et al*., 2012; Wolff *et al*., 2007). This lack of consistency in the literature was also observed for birth outcomes in the recent chlorpyrifos HHRA (USEPA, 2014) which notes that researchers from the three US birth cohort studies also investigated the possible role of prenatal OP exposure and fetal growth. These results were not consistent across these cohorts. Authors with Columbia Center for Children's Environmental Health (CCCEH) Mothers and Newborn Study observed evidence of an inverse association, *i.e.*, increasing cord blood chlorpyrifos was associated with decreased measures of birth weight and length, while authors with the Mt. Sinai and CHAMACOS cohorts reported either no association, or evidence of a *positive* relationship, respectively (Berkowitz *et al*., 2004; Eskenazi *et al*., 2004; Whyatt *et al*., 2004). Inconsistent results may be due to differences across study groups in exposure profiles as well as dissimilar methods of prenatal OP exposure assessment (Needham, 2005). Given the lack of consistency among cohorts for the fetal growth metrics, the proposed link between fetal growth and OP exposure is tenuous. Therefore, consistent with previous evaluations for chlorpyrifos, EPA is focusing the remainder of this document on neurodevelopmental outcomes. Although the agency is not evaluating these birth outcome studies further at this time, the agency will continue to monitor the scientific literature for advances in this line of research.

### 2.2.8 Summary of Epidemiology Studies from 2015 Literature Review Focusing on Neurodevelopmental Outcomes

From an initial total of 23 high or moderate quality studies, with the exclusion of six studies for insufficient exposure assessment, four studies with measurable AChE inhibition and potential cholinergic toxicity which are outside the scope of this analysis, one study from an illegal use of a highly potent OP (MP) where cholinergic toxicity cannot be ruled out, four studies with birth

outcome as the only assessment, and one study assessing only PON1 genotype expression and neurobehavioral outcomes (Fortenberry et al, 2014a), a total of seven studies focusing on neurodevelopmental outcomes remain to be evaluated.

Seven total studies focused on OP exposure and either neurodevelopmental or neurobehavioral outcomes, with three high quality (Bouchard *et al.*, 2010; Fortenberry *et al.*, 2014; Furlong *et al.*, 2014) and four moderate quality studies (Guodong *et al.*, 2012; Oulhote and Bouchard, 2013; Zhang *et al.*, 2014; Shelton *et al.*, 2014). The results of these studies are summarized below. With the exception of Shelton *et al.* (2014), all of these studies used biomarker measures for their exposure assessment, including child or maternal urinary DAPs (Bouchard *et al.*, 2010; Furlong *et al.*, 2014; Guodong *et al.*, 2012; Oulhote and Bouchard, 2013; Zhang *et al.*, 2014), child urinary DMP and DEPs (Oulhote and Bouchard, 2013; Zhang *et al.*, 2014; Guodong *et al.*, 2012), or maternal TCPy (Fortenberry *et al.*, 2014). Shelton *et al.* (2014) used pesticide use data from the California Department of Pesticide Regulation and geospatial methods to map the specific pesticide use pattern to the participant residence. The study populations ranged from national in scope (Bouchard *et al.*, 2010; Oulhote and Bouchard, 2013) to mainly urban (Fortenberry *et al.*, 2014; Fortenberry *et al.*, 2014a; Furlong *et al.*, 2014; Guodong *et al.*, 2012; Zhang *et al.*, 2014; Shelton *et al.*, 2014).

In order to summarize the results from these studies, we have separated them into three groups based on the outcomes assessed — two Chinese studies focusing on more generic neurodevelopmental outcomes (Zhang *et al.*, 2014; Guodong *et al.*, 2012); two studies focusing on social responsiveness outcomes or autism spectrum disorders (Furlong *et al.*, 2014; Shelton *et al.*, 2014); and three studies focused on ADHD, behavioral problems, and attention problems (Oulhote and Bouchard, 2013; Bouchard *et al.*, 2010; Fortenberry *et al.*, 2014).

In the first of two Chinese studies focusing on generic neurodevelopmental outcomes, Zhang *et al.* (2014) investigated prenatal exposure to OPs and neurobehavioral development of neonates in a birth cohort study in Shenyang, China. The authors reported that consistent statistically significant associations were observed between all of the quantified urinary biomarkers of prenatal OP pesticide exposure and neonatal neurodevelopment deficits assessed 3 days after birth. A 10-fold increase in total DAPs concentration was associated with an average decrease in Neonatal Behavioral Neurological Assessment (NBNA) summary scores of 1.78 points (95% CI: -2.12 to -1.45). No evidence of departure from linearity of the exposure-response relationship between maternal DAP concentrations and NBNA scores was observed. In the second Chinese study focusing on neurodevelopmental outcomes, Guodong *et al.* (2012) conducted a cross-sectional study, and did not identify any statistically significant associations between the children's urinary DAP metabolite levels and any of the DQ (Developmental quotients) scores. The authors mentioned that their results should be interpreted with caution since OP exposure was quantified in single spot urine sample from children, and should be followed up with a longitudinal study with repeated measurement of exposure levels in urine samples.

Two studies focused on impaired social responsiveness, autism spectrum disorders, or developmental delays (Furlong *et al.*, 2014; Shelton *et al.*, 2014). Among 7-9 years old children in the Mount Sinai Cohort (Furlong *et al*. 2014), there was no overall statistically significant association between maternal third trimester urinary DAP metabolite levels and reciprocal social responsiveness (a measure linked to many neuropsychiatric conditions that involve impaired social functioning (Constantino and Gruber 2005)). However, some evidence of modification of the association between prenatal OP pesticide exposure and impaired social responsiveness in early childhood was observed by both race/ethnicity and child sex, with an association between DEP and poorer social responsiveness observed among black participants and boys. No association was observed among whites or Hispanics, among girls, or for DAP or DMP biomarker levels.

Shelton *et al.* (2014), investigated autism spectrum disorders (ASD) and developmental delay (DD) in relation to gestational residential proximity to agricultural pesticide applications utilizing the California population-based Childhood Autism Risks from Genetics and Environment (CHARGE) study. The investigators reported that children with ASD were 60% more likely to have OPs applied near the home [1.25 km distance; adjusted OR (aOR) = 1.60; 95% CI: 1.02–2.51] than mothers of normally developing children. They added that as the buffer distance grew larger, these associations became lesser, indicating an exposure-response effect. The authors also mentioned that each 100-lb (45.4 kg) increase in the amount applied over the course of pregnancy (within 1.5 km of the home) was associated with a 14% higher prevalence of ASD (aOR = 1.14; 95% CI: 1.0, 1.32), but no association was identified with DD.

A total of three studies focused on OP exposure and behavioral, memory, or attention/ADHD outcomes (Oulhote and Bouchard, 2013; Bouchard *et al.*, 2010; Fortenberry *et al*., 2014). In a national cross-sectional study of Canadian children 2007-2009 data for age 6-11 years (Oulhote and Bouchard, 2013), there were no overall statistically significant associations observed between child urinary DAP, DMP, or DEP metabolite levels and parentally reported behavioral problems. In contrast, Bouchard *et al*. (2010), looking at U.S. children age 8-15 years in the 2000-2004 National Health and Nutrition Examination Survey (NHANES),[5] observed a positive association between attention and behavior problems and DAPs and DMPs, but not DEPs. For example, even after controlling for potential confounders such as sex, age, ethnicity, and creatinine concentration, they found that a 10-fold increase in DMAP concentration was associated with a 55 to 72% increased odds of ADHD.

Fortenberry *et al*. (2014) evaluated the relationship between pesticide exposure and ADHD in school age Mexican children, recruiting 187 mother-child pairs from a prospective birth cohort, ELEMENT (Early Life Exposures in Mexico to Environmental Toxicants). The authors reported that, there were no statistically significant associations between tertiles of maternal third trimester urinary TCPy and measures of attention and hyperactivity in children. However, there was suggestive evidence for increases in the ADHD index in relation to TCPy tertiles among boys

[5] http://www.cdc.gov/nchs/nhanes.htm

(the highest TCPy tertile was associated with an ADHD index score that was 5.55 points higher than children in the lowest tertile; p-value = 0.06).

### 2.2.9 Summary of Three US Children's Cohort Studies (CCCEH, CHAMACOS, Mt. Sinai): Focusing on Neurodevelopmental Outcomes Evaluated in 2012/2014

Detailed summaries and evaluation, associated strengths and limitations, and accompanying detailed evidence table for the CCCEH, CHAMACOS, and Mt. Sinai cohorts can be found in the white paper for the 2012 SAP review and the 2014 chlorpyrifos revised HHRA. Limited summary information is provided here for comparison with the studies identified in the 2015 literature review discussed above.

#### 2.2.9.1 Overview

In the chlorpyrifos revised HHRA, EPA included epidemiologic research results from three prospective birth cohort studies. These include: 1) The Mothers and Newborn Study of North Manhattan and South Bronx performed by the Columbia Children's Center for Environmental Health (CCCEH) at Columbia University; 2) the Mt. Sinai Inner-City Toxicants, Child Growth and Development Study or the "Mt. Sinai Child Growth and Development Study;" and 3) the Center for Health Assessment of Mothers and Children of Salinas Valley (CHAMACOS) conducted by researchers at University of California Berkeley. In these epidemiology studies, mother-infant pairs were recruited for the purpose of studying the potential health effects of environmental exposures during pregnancy on subsequent child development. Importantly, each of these cohorts evaluated the association between prenatal chlorpyrifos or OP exposure with adverse neurodevelopmental outcomes in children through age 7 years.

These studies reflect different types of exposed groups in the total population which strengthens the weight of the evidence considerations regarding this stream of information. The CCCEH Mother's and Newborn study and the Mt. Sinai Child Growth and Development study participants were likely exposed to OPs through the diet and through residential use of the pesticide for indoor pest control. In the residential setting, study populations were most likely exposed through indoor residential use of the pesticide during the study time period and additionally exposed to OPs via the oral route through ingesting residues in the diet and from hand-to-mouth contact with in-home surfaces, as well as possible dermal or inhalation exposure through contact with treated areas in the home environment (Berkowitz *et al.*, 2003; Whyatt *et al.*, 2003; Whyatt *et al.*, 2009; Whyatt *et al.*, 2007). In contrast, CHAMACOS cohort participants were employed as farm laborers or were residing in homes with farm laborers. The CHAMACOS study participants likely experienced exposure to OPs through the diet and from occupational exposure (primarily inhalation and dermal routes), as well as probable indirect take-home exposures (the "tracking in" of pesticide residues through shoes and clothing, augmented by poor hygiene practices) (Bradman *et al.*, 2007). In each of the three US children's health cohorts, the biological measurements in these cohorts were comparable to the general population NHANES. EPA has considered the strengths and limitations of these

studies, and believes that random or systematic errors in the design, conduct or analysis of these studies were unlikely to fully explain observed positive associations between *in utero* OP exposure and adverse neurodevelopmental effects observed at birth and through childhood (age 7 years). EPA believes these are strong studies which support a conclusion that OPs likely played a role in these outcomes.

### 2.2.9.2 Review of Study Design and Research Methods

These cohort studies each enrolled pregnant women during roughly the same time period, measured both environmental exposure to the pesticide during pregnancy and also measured biomarkers representing internal dose during pregnancy and at delivery, and prospectively assessed associations in their newborns and young children through age 7 years. Each study includes several hundred (approximately 100-400) mother-infant pairs; these sample sizes are sufficient to perform statistically valid analyses. Investigators from each study cohort utilized a similarly strong study design (prospective birth cohort); measured pesticide exposure using several different methods including environmental indicators as well as specific and non-specific biomarkers of OPs; ascertained developmental outcomes using validated assessment tools well-established in both clinical and research settings; and, measured, analyzed, selected and statistically adjusted for potentially confounding variables including socio-economic status and other environmental exposures using reasonable and appropriate methods. Limitations exist as well. These studies utilized a one-time measure (or the average of two measures) of chlorpyrifos or OP exposure to assess prenatal pesticide exposure throughout the gestational period, were unable to assess the influence of mixtures (co-occurring exposures in the relevant biological time window), and reflect a small sample size to fully evaluate the effect of more than one simultaneous exposure on neurodevelopment, *i.e.*, evidence of effect modification.

As noted, two major uncertainties in environmental epidemiology studies are the accurate and reliable measurement of exposure and potential confounding variables such as the influence of mixtures. The researchers with each of the three cohorts have provided supplemental methodological research to address these areas to the extent possible. Across the three children's health cohorts, study authors measured biomarkers of OP exposure. There is uncertainty as to the extent measurement of non-specific metabolites of OP or chlorpyrifos accurately reflects OP exposure; CCCEH and Mt. Sinai studies do not estimate post-natal exposure to chlorpyrifos among child participants, therefore the influence of early life and childhood OP exposure is unaccounted for in these analyses. The CHAMACOS cohort measured urinary levels of DAPs in young children and did not observe negative significant associations in relation to neurodevelopment from post-natal exposure (Eskenazi *et al.*, 2007). The CHAMACOS cohort investigators also measured AChE and butyl ChE as supplemental indicators of OP exposure.

Potential confounding bias is another major uncertainty within environmental epidemiology studies. Confounding variables, exposures that could be related to OP exposure and neurodevelopmental outcomes such as blood lead, may result in an incorrect epidemiological risk estimate. Across these cohort studies, investigators collected relevant information

concerning demographic characteristics and other environmental exposures, and were, to the extent possible with the existing information, able to effectively hold constant the influence of these other variables when estimating the association between prenatal chlorpyrifos and adverse neurodevelopmental outcomes. Control of these variables is important to reduce the chances of a false positive study result. Overall, statistical analyses were judged to be appropriate and reasonable (not overly large number of statistical model variables) to the research question by EPA and expert Panel reviews (FIFRA SAP 2008 and 2012).

Researchers with both the Mt. Sinai and CHAMACOS cohorts evaluated neonatal neurological functioning in association with prenatal OP exposure; CCCEH did not conduct these measurements. To measure indices of abnormal neonatal behavior and/or neurological integrity authors used outcome measures derived from the Brazelton Neonatal Behavioral Assessment Scale (BNBAS), a neurological assessment of 28 behavioral items and 18 primitive reflexes. This tool was administered to infants 2-5 days post-partum by trained neonatologists in the hospital setting using similar environmental conditions. The authors with both study groups observed an increased number of abnormal reflexes in relation to increasing measures of OP exposure (Engel *et al.*, 2007; Young *et al.*, 2005). Among the other 27 measures in the BNBAS, neither study group reported evidence of any other positive associations. The authors also observed evidence of potential effect modification by PON1 activity level in the relation between DAPs and neonatal neurodevelopment in which infants of mothers who are slower metabolizers have greater risk of abnormal reflexes (Young *et al*. 2005; Engel *et al*. 2007). However, EPA notes these studies are likely under-powered to make a statistically robust estimate of this statistical interaction.

Researchers across the three children's health cohorts utilized the Bayley Scales of Infant Development II (BSID-II) to generate a Mental Development Index (MDI) and a Psychomotor Development Index (PDI) to assess neurodevelopment in early childhood. In the CCCEH Mothers and Newborn study, Rauh *et al*. (2006) investigated MDI and PDI at 12, 24, and 36 months of age. Children were categorized as having either high (>6.17pg/g) or low ($\leq$6.17pg/g) prenatal chlorpyrifos exposure, using categories informed by results of the previous study on birth characteristics (Whyatt *et al*., 2004). Authors reported that the difference in MDI scores was "marginally significant" (p = 0.06) between the "high" and "low" exposed groups; the high exposed group scoring an average of 3.3 points lower than the low exposed (Rauh *et al*., 2006). Regarding the PDI score (motor skills), none of the 12 or 24 month PDI scores showed significant effects, but the 36 month score was significantly related to chlorpyrifos exposure. Researchers noted that the effects were most pronounced at the 36 month testing period. Within the 36 month testing period, the likelihood of highly exposed children developing mental delays were significantly greater (MDI: 2.4 times greater (95% CI: 1.12-5.08, p = 0.02) and PDI: 4.9 times greater (95% CI: 1.78-13.72; p = 0.002)) than those with lower prenatal exposure (Rauh *et al*., 2006). Within the Mt. Sinai study, authors administered the BSID-II to participating children at 12 and 24 months and observed that prenatal total DAP metabolite level was associated with a decrement in mental development at 12 months among blacks and Hispanic children; however, these associations either attenuated or were non-existent at the 24-month visit (Engel *et al*., 2011). In the CHAMACOS cohort, Eskenazi *et al*. (2007) observed

that prenatal DAP levels were adversely associated with MDI, and at 24 months of age these associations reached statistical significance. In this study, neither prenatal DAPs nor maternal TCPy were associated with PDI (motor skills), nor did authors observe evidence of different risk by PON1 status (Eskenazi *et al.*, 2010).

With respect to the findings related to the autism spectrum, from CCCEH, Rauh *et al.* (2006) reported a large odds ratio for pervasive developmental disorder (PDD) (OR=5.39; 95% CI: 1.21-24.11) when comparing high to low chlorpyrifos exposure groups. As described above, among 7-9 years old children in the Mount Sinai Cohort (Furlong *et al.* 2014), there was no overall statistically significant association between maternal third trimester urinary DAP metabolite levels and reciprocal social responsiveness. However, some evidence of modification of the association between prenatal OP pesticide exposure and impaired social responsiveness in early childhood was observed by both race/ethnicity and child sex, with an association between DEAP and poorer social responsiveness observed among black participants and boys. No association was observed among whites or Hispanics, among girls, or for DAP or DMAP biomarker levels. In the CHAMACOS cohort, Eskenazi *et al.* (2010) reported non-significant, but suggestive, increased odds of PDD of 2.0 (0.8 to 5.1; p=0.14), whereas Eskenazi et al (2007) reported a statistically significant association between total DAP exposure and increased odds of PDD.

With respect to attention problems, Rauh *et al.* (2006) also investigated 36-month child behavior checklist (CBCL) (behavioral) scores. Significant differences were observed between the high and low chlorpyrifos exposure groups in the general category of attention-problems (p=0.010), and in the more specific DSM-IV scale for ADHD problems (p = 0.018). The CHAMACOS cohort also investigated attention problems in early childhood using three different assessment tools: maternal report of child behavior at 3.5 and 5 years of age; direct assessment of the child at 3.5 and 5 years; and by a psychometrician's report of the behavior of the child during testing at 5 years. In this study population, higher concentrations of OP metabolites in the urine of pregnant women were associated with increased odds of attention problems and poorer attention scores in their children at age 5 years (Eskenazi *et al.*, 2007).

To measure intelligence among school aged children, authors from each of the three children's health cohorts used the Wechsler Intelligence Scale for Children, 4th edition (WISC-IV). The instrument measures four areas of mental functioning: the Verbal Comprehension Index, the Perceptual Reasoning Index, the Working Memory Index, and the Processing Speed Index. A Full-Scale IQ score combines the four composite indices. WISC-IV scores are standardized against U.S. population-based norms for English and Spanish-speaking children. In the CCCEH Mothers and Newborn Study, Rauh *et al.* (2011) evaluated the relationship between prenatal chlorpyrifos exposure and neurodevelopment among 265 of the cohort participants who had reached the age of 7 years and had a complete set of data including prenatal maternal interview data, prenatal chlorpyrifos marker levels from maternal and/or cord blood samples at delivery, postnatal covariates, and neurodevelopmental outcome data (Rauh *et al.*, 2011). While models were developed using continuous measures of both prenatal chlorpyrifos exposure and Wechsler scores, for ease of interpretation, investigators reported that for each

standard deviation increase in exposure (4.61 pg/g) there is a 1.4% reduction in Full-Scale IQ and a 2.8% reduction in Working Memory. In the Mt. Sinai study, prenatal maternal DEP urinary metabolite concentrations were associated with slight decrements in Full Scale Intelligence Quotient (FSIQ), Perceptual Reasoning, and Working Memory between the ages of 6 and 9 years, and difference in intelligence measures by putative PON1 status were also noted (Engel *et al.*, 2011). Similarly, in the CHAMACOS cohort, Bouchard *et al*. (2011) observed evidence of an association between prenatal exposures to OPs as measured by urinary DAP (total DAP, DEP, and DMP) metabolites in women during pregnancy, and decreased cognitive functioning in children at age 7. In this study, children in the highest quintile of maternal DAP concentrations had a statistically significant 7 point difference in IQ points compared with those in the lowest quintile.

To ascertain whether observed differences in neurodevelopment after prenatal chlorpyrifos exposure may be explained by differences in brain morphology between exposed groups, investigators compared MRI brain images between high and low chlorpyrifos exposed child study participants (Rauh *et al.*, 2012). Authors determined there were distinct morphological differences in brain areas associated with these neurodevelopmental outcomes. The pilot study included 40 child participants due to strict inclusion and exclusion criteria, and the high cost of performing the imaging studies on each child. EPA convened a Federal Panel of experts to perform a written peer-review of this study.[6] The Federal Panel concurred with the authors' conclusions in general; however the Federal Panel also noted that significantly greater and more sophisticated MRI imaging studies would be needed to link the morphological changes indicated in this study with specific functional outcomes noted in the CCCEH IQ study. Therefore, while generally supportive of the epidemiologic findings, additional study is needed to make specific links with areas of brain development change.

In sum, across these three children's environmental health studies, authors consistently identified associations with neurodevelopmental outcomes in relation to OP exposure. There is evidence of delays in mental development in infants (24-36 months), attention problems and autism spectrum disorder in early childhood, and intelligence decrements in school age children who were exposed to chlorpyrifos or OPs during gestation. Investigators reported strong measures of statistical association across several of these evaluations (odds ratios 2-4 fold increased in some instances), and observed evidence of exposures-response trends in some instances, *e.g.*, intelligence measures.

[6] http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2008-0850-0170.

## Table 2.2.9 Detailed Summary Tables of Children's Environmental Health Epidemiology Studies (extracted from USEPA, 2014)

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| Article 1: Whyatt *et al.* (2004) Columbia U. (N=314) | Birth length, Birth weight, head circumference | 4 cord plasma chlorpyrifos exposure groups and 4 chlorpyrifos and diazinon exposure groups. Chlorpyrifos only categories, Group 1: levels below LOD (32% of participants); Group 2: lowest 1/3 of detectable levels (20 %); Group 3: middle 1/3 (24%), Group 4: highest 1/3 (25%.) Chlorpyrifos and diazinon together: Group 1: 26%, Group 2: 22 %, Group 3: 26%, Group 4: 26%. | Gestational age, maternal pre-pregnancy weight, maternal net pregnancy weight gain, gender of newborn, parity, race/ethnicity, ETS in home, season, cesarean section | For each log unit increase in cord plasma chlorpyrifos levels, birth weight decreased by 42.6 g (95% CI: –81.8 to –3.8) and birth length decreased by 0.24 cm (95% CI: –0.47 to –0.01). Birth weight averaged 186.3 g less (95% CI: –375.2 to –45.5) among newborns with the highest compared with lowest 26% of exposure levels (p = 0.01). | Associations between birth weight and length and cord plasma chlorpyrifos were statistically significant ($p \leq 0.007$) among newborns born before the January 2001 policy change. Among newborns born after January 2001, exposure levels were substantially lower, and no associations with fetal growth outcomes were observed ($p > 0.8$). | Strengths: prospective nature of the study; direct measurement of chlorpyrifos in cord blood and personal air samples, rather than non-specific markers of organophosphate pesticide exposure; consideration of other pesticides and environmental contaminants as covariates in the multivariate models. Limitations: single exposure sampling period; the authors did not present nor discuss regression diagnostics to assess the degree to which their models met or violated the assumptions implicit in linear models. |
| Article 2: Berkowitz *et al.*(2004) Mt. Sinai (N=404) | Birth length, birth weight, head circumference, gestational age | LOD: 11 ug/L (57% <LOD TCPy) | Race/ethnicity, infant sex, and gestational age. The authors also controlled for birth weight or birth length in their assessment of head circumference and pesticide exposure. | Mean levels of birth weight, length, head circumference, and gestational age did not differ between those with urinary pesticide metabolite levels below and above the level of detection.<br><br>Similarly, no statistically significant associations were observed between reported pesticide exposure and mean indices of fetal growth and | PON1 activity also predictor of smaller head circumference; creatinine corrected | Very well conducted study with numerous strengths and very few weaknesses.<br><br>The questionnaire-based pesticide exposure questions are subject to imperfect recall. Errors would, on average, attenuate associations between these exposure metrics and fetal development.<br><br>Recall-based exposure assessments were fortified |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | gestational age. | | by objective measures of pesticides/pesticide metabolites.<br><br>A metabolite specific for chlorpyrifos (TCPy) was assessed.<br><br>Statistical analysis was appropriate.<br><br>Observed mean reductions in the outcome parameters appear to be small in magnitude and may be of little clinical significance.<br><br>Assessment and control for confounding were appropriate. However, confounding by unmeasured (and mismeasured) risk factors for abnormal growth that are related to pesticide exposure would bias the results in this study. Such factors may be related to socioeconomic status of the study participants, a cofactor which is difficult to define, no less measure in an epidemiologic study.<br><br>Limited external validity (generalizability) due to the particular study population recruited and the numerous exclusion criteria applied. |
| Article 3: Eskenazi *et al.* (2004) CHAMACOS | Birth length, birth weight, head circumference, | Total DAPs: median 136 nmol/L (range: | Gestational age, gestational age squared, maternal age, pregnancy weight gain, week of | Decreases in gestational duration associated with two measures of in utero | Maternal urine collection averaged weeks 14, 26, not creatinine-corrected | Strengths in the study design include the longitudinal design, the use |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| (N=488) | Gestation Length, Ponderal index | 10–6,854); DEP: median 22 nmol/L (range: 2–680 nmol/L); TCPy: median 3.3 nmol/L (range: 0.2–56.1nmol/L) (76% >LOD) | initiating prenatal care, parity, infant sex, mother's country of birth, body mass index, family income, poverty level, smoking, alcohol, illicit drug use, environmental tobacco smoke, caffeine, history of low birth weight, and history of pre-term delivery. | pesticide exposure: levels of metabolites of dimethyl phosphate pesticide compounds and whole blood ChE. | | of multiple exposure biomarkers, including quantification of non-specific (DAPs), chlorpyrifos-specific (TCPy) metabolites, and other environmental co-exposures. A reasonable set of exclusion criteria were applied.The selection of the CHAMACOS population, which consists mostly of children from low-income families, served to increase the relative statistical efficiency of the study, as this population is at high risk of neurodevelopmental deficits, compared to the general population.The statistical analysis used to assess the associations between the markers of exposure and neurodevelopment were appropriate.Errors in the assignment of exposure in this prospective study will likely have resulted in attenuation of observed associations. |
| Article 4: Harley *et al*. (2011) CHAMACOS (N=329) | Birth length, birth weight, head circumference, gestational age | The geometric mean for the DAP concentrations during pregnancy (for the average of the two sampling periods) was 146 nmol/L (95% CI: 133, 160); of this, a larger proportion | Maternal intelligence (Peabody Picture Vocabulary Test (PPVT)), measures of how stimulating the environment is, and known or suspected neurotoxins were measured prenatally. To measure the quality and extent of stimulation available to a child in the home environment, the Infant-Toddler HOME (Home Observation for Measurement | The authors observed evidence of an association between prenatal exposure to OP pesticides as measured by urinary DAP metabolites in women during pregnancy, is associated with decreased cognitive functioning in children at age 7. | Infants whose PON1 genotype and enzyme activity levels suggested that they might be more susceptible tothe effects of OP pesticide exposure had decreased fetal growth and length of gestation. PON1 may be a contributing factor to preterm or low birth weight birth. | This study has many strengths, the longitudinal design, the measurement of urinary DAP at multiple time points and following children to age seven when tests of cognitive function are reportedly more reliable. The authors were able to adjust for or consider many factors |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | was DMP metabolites (GM = 109 nmol/L; 95% CI =98, 120) than DEPs (GM= 23 nmol/L; 95% CI = 21, 25). Allele frequencies: PON1 192 Q allele= 50%; PON1 -108 T allele= 46%.Mean arylesterase activity: For infants: 33.6 U/mL (SD = 16) For mothers: 136.6 U/mL (SD = 44). Mean paraoxonase activity: For infants: 256.6 U/L (SD = 165); For mothers: 989.0 U/L (SD = 616). | of the Environment) inventory was completed at the 6-month, 1, 2, 3.5, 5, and 7 year visits; known or suspected neurotoxicants, polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs), p,p'-dichlorodiphenyltrichlorethane (DDT), p,p'-dichlorodiphenyltrichlorethylene (DDE), and lead. | | | related to cognitive function, such as prenatal exposure to other environmental agents, socioeconomic indicators, maternal intelligence and education, and child stimulation. The cohort had a relatively homogenous socioeconomic profile, reducing the potential for uncontrolled confounding. |
| Article 5: Engel *et al*. (2007) Mt. Sinai (N=311) | Brazelton Neonatal Behavioral Assessment Scale (BNBAS), primitive reflexes (neurological integrity) measured before hospital discharge. | DEP: 24.7 nmol/L; Total DAP: 82 nmol/L | Maternal age, race, marital status, education, cesarean delivery, delivery anesthesia, infant age at examination, infant gender, infant jaundice, smoking (yes/no), alcohol consumption, caffeine consumption, illicit drug use during pregnancy, and the examiner. | No adverse associations were found for DAPs and any measured behavior. Relative to the first quartile, quartiles 2–4 of total DEPs, DMPs, and DAPs were associated with an increased proportion of abnormal reflexes, although the associations did not increase monotonically and varied in their strength and precision. | Used non-specific biomarker DEP/DAP | This was a well conducted prospective study conducted in a young, predominantly minority population. The study design, analytic approach, and statistical analyses were appropriate. Pesticide metabolites evaluated are not specific for chlorpyrifos. The BNBAS was administered before hospital discharge only on a subset of children in the cohort (n =311/404). Factors related to weekend delivery (e.g., fewer inductions) would be |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | underrepresented among the tested subjects, and may induce bias, reduce the degree of precision with which associations were estimated, and limit the generalizability of the study findings.The statistical analysis was largely appropriate. Imputing of missing data may affect precision of association estimates and result in attenuated effect estimates as a result of exposure measurement error. Assessment and control for confounding were appropriate. However, confounding by unmeasured (and mis-measured) risk factors for abnormal growth that are related to pesticide exposure would bias the results in this study. Such factors may be related to socioeconomic status of the study participants, a cofactor which is difficult to define, no less measure in an epidemiologic study. |
| Article 6: Young *et al*. (2005) CHAMACOS (N=381) | Neurodevelopment, Brazelton Neonatal Behavioral AssessmentScale (BNBAS), abnormal reflexes | DAP (average during pregnancy): median 222nmol/L (range: 7–21,867 nmol/L); DEP (average during pregnancy): median 21 nmol/L (range: 2–680 | Maternal age, BMI, any smoking/alcohol/drug use during pregnancy, gestational age at which prenatal care was initiated, total number of prenatal care visits, mean pregnancy blood pressure, parity, method of delivery, general anesthesia used during delivery, breastfeeding initiated after delivery, poverty level, | Among the >3 day old infants, increasing average prenatal urinary metabolite levels were associated with both an increase in number of abnormal reflexes (total DAP: adjusted beta = 0.53, 95% CI = 0.23, 0.82; dimethyls: adjusted beta = 0.41, 95% CI = 0.12, 0.69; diethyls: adjusted beta = | Associations seen pre-natal OP, not post-natal OP exposure, Maternal urine collection averaged weeks 14, 26 | Strengths: Longitudinal design, measurement and consideration of many confounders, the prenatal exposure measures were, with some exceptions, the average of two measurements, and thus may better reflect chronic exposure during the pregnancy |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | nmol/L) | infant sex, age in days at BNBAS, minutes since last feed at BNBAS, and BNBAS examiner. | 0.37, 95% CI = 0.09, 0.64), and the proportion of infants with more than three abnormal reflexes (total DAP: adjusted OR = 4.9, 95% CI = 1.5, 16.1; dimethyls: adjusted OR = 3.2, 95% CI = 1.1, 9.8; diethyls: adjusted OR = 3.4, 95% CI = 1.2, 9.9). | | Weaknesses: Potential for residual confounding, potential for exposure misclassification, potential selection bias (not all children had the outcome measures), study population not generalizable to the whole US. |
| Article 7: Rauh *et al*. (2006) Columbia U. (N=254) | Neurodevelopment: The Bayley Scales of Infant Development II (BSID-II), Mental Development Index (MDI) and Psychomotor Development Index (PDI) at 12, 24, and 36 months of age. • Behavior: Child Behavior Checklist (CBCL) at 12, 24, and 36 months. • Quality of the child-care environment: The Home Observation for Measurement of the Environment (HOME) | Exposure levels were categorized as low (≤ 6.17 pg/g) or high (>6.17 pg/g) | Data were collected regarding lead exposure, demographics, education and occupational history, income, active and passive smoking, alcohol and drug use during pregnancy, and residential pesticide use. Final models included prenatal environmental tobacco smoke (ETS) exposure, gender, ethnicity, gestational age at birth, quality of home care-taking environment, maternal education, and maternal IQ. | At the 36 month milestone, the likelihood of highly exposed children developing mental delays were 2.4 times greater (95% CI: 1.12-5.08, p = 0.02) and motor delays were 4.9 greater (95% CI: 1.78-13.72; p = 0.002) than those with lower prenatal exposure. The GLM analysis for PDI scores showed a significant effect of chlorpyrifos exposure over time with an estimated deficit of approximately 7 points by age 36 months (p = 0.01). | The authors summarize three main findings: 1) by age 3, the more highly exposed children demonstrated mental and motor delays; 2) the observed developmental trajectories for PDI and MDI scores confirmed that the adverse impact on cognitive and motor development increased over time; and 3) by age 3, highly exposed children were more likely to demonstrate clinically significant attention problems. | • Only 53% of the children reached the three year milestone with study data collected. It is unclear what percentage of these children did not survive, were lost to follow-up, or too sick to participate. • Reliance on a single exposure level (prenatal/cord blood.) • No control for exposure over the subsequent 3 years • Creation of a dichotomous exposure variable brings limitations due to the amount of within-group variation. • Limitations of the sensitivity and predictive validity of the developmental tests, especially among children less than 3 years of age. • No discussion of whether this 7-point deficit is clinically relevant. • Due to the pervasive, non-specific nature of neurological effects, it is difficult to attribute causality. |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| Article 8: Lovasi *et al*. 2010 Columbia U. (N=266) | Bayley scores (MDI/PDI) 12 months, 24 months, 36 months | N/A | Neighborhood characteristics: The percentage of housing units without complete plumbing, the percentage of vacant housing units, the percentage of residents below the federal poverty line, the percentage of residents older than 25 years of age who completed high school, the percentage of households receiving public assistance, the percentage of housing units with one or more residents per room, racial composition, the percentage of residents born outside the United States, the percentage of Spanish-speaking residents, and the percentage of residents who were linguistically isolated | Neighborhood characteristics did not confound the observed association between chlorpyrifos levels and cognitive development. | Hierarchal regression analysis of potential confounding by SES | Direct measurement of chlorpyrifos.The statistical analyses were generally appropriate. Missing data on covariates were estimated using multiple imputation, and the variance estimates presented appropriately reflect the degree of uncertainty caused by missing covariate data. Robust standard errors were used.The setting of the investigation in a sample drawn from low-income African American and Dominican communities is both a strength (increases the power, restriction of confounders) and a limitation of the study (reduced generalizability). |
| Article 9: Engel *et al*. (2011) Mt. Sinai (N=276) | Bayley scores (MDI/PDI) at 12 months, 24 months. | DEP: 24.7 nmol/L; Total DAP: 82 nmol/L (same as Engel 2007) | Maternal age, race/ethnicity, marital status, education, breast-feeding, child sex, alcohol, smoking, or drug use during pregnancy, maternal IQ, a score based on assessment of the home environment (HOME), season of urine collection, language spoken in the home, age at testing, examiner, and urinary creatinine level. | An observed association between prenatal total dialkylphosphate metabolite level and a decrement in mental development at 12 months among blacks and Hispanics. | Used non-specific biomarker DEP/DAP; some evidence of effect modification by PON1 genotype | Limitations include use of non-specific markers of chlorpyrifos pesticide exposure (DAPs), use of only a single (third-trimester) urine sample, and the large proportion of loss to follow-up. Statistical analysis was appropriate. Imputing of missing data may affect precision of association estimates and result in attenuated effect estimates as a result of exposure measurement error, although these are offset by the further categorization of the exposure levels (at the |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | median). However, binning of exposure levels reduces precision, relative to a continuously distributed measure of exposure. Assessment and control for confounding were appropriate. However, confounding by unmeasured (and mismeasured) risk factors for abnormal growth that are related to pesticide exposure would bias the results in this study. Such factors may be related to socioeconomic status of the study participants, a cofactor which is difficult to define, no less measure in an epidemiologic study. |
| Article 10: Eskenazi *et al*. (2007) CHAMACOS (N=372) | Neurodevelopment, Bayley Index (MDI, PDI), Maternal behavior checklist at 6, 12, and 24 months | DEP: geom. Mean in mother 18.1 nmol/L (95% CI = 16.7–19.7); DEP geometric mean in child at 24 months 10.5 nmol/L (95% CI =8.8–12.6); TCPy median 3.54 ug/l | Psychometrician, location of assessment, exact age at assessment, sex, breast-feeding duration (months), HOME score, and household income, parity, maternal PPVT, maternal age, education, depressive symptoms, active/passive smoking exposure during pregnancy, regular alcohol use during pregnancy, marital status, father's presence in home, housing density, maternal work status, ≥ 15 hours out-of-home childcare/week, birth weight, gestational age, abnormal reflexes, PCBs, lead, DDT, β-hexachlorocyclohexane, and hexa-chlorobenzene | DAP metabolite levels during pregnancy, particularly from dimethyl phosphate pesticides, may be negatively associated at 24 months with mental development (MDI) on the Bayley Scales and an increase in risk of maternally reported PDD. | No strong associations identified with DE or TCPy, Maternal urine collection averaged weeks 14, 26 | Strengths: Longitudinal design, measurement and consideration of many confounders (including other environmental chemicals); the prenatal exposure measures were, with some exceptions, the average of two measurements, and thus may better reflect chronic exposure during the pregnancy<br><br>Weaknesses: Potential for residual confounding, potential for exposure misclassification, potential selection bias (not all children had the outcome measures), study population not |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | generalizable to the whole US. |
| Article 11: Eskenazi *et al*. (2010) CHAMACOS (N=371) | Neurodevelopment, Bayley Index (MDI, PDI), Maternal behavior checklist at 6, 12, and 24 months, PON1 gene and enzyme levels | The geometric mean for the DAP concentrations during pregnancy (for the average of the two sampling periods) was 146 nmol/L (95% CI: 133, 160); of this, a larger proportion was DMP metabolites (GM = 109 nmol/L; 95% CI =98, 120) than DEPs (GM= 23 nmol/L; 95% CI = 21, 25). Allele frequencies: PON1 192 Q allele= 50%; PON1 -108 T allele= 46%.Mean arylesterase activity: For infants: 33.6 U/mL (SD = 16) For mothers: 136.6 U/mL (SD = 44). Mean paraoxonase activity: For infants: 256.6 U/L (SD = 165); For mothers: 989.0 U/L (SD = 616). | Psychometrician, location of assessment, exact age at assessment, sex, breast-feeding duration (months), HOME score, and household income, parity, maternal PPVT, maternal age, education, depressive symptoms, active/passive smoking exposure during pregnancy, regular alcohol use during pregnancy, marital status, father's presence in home, housing density, maternal work status, ≥ 15 hours out-of-home childcare/week, birth weight, gestational age, abnormal reflexes, PCBs, lead, DDT, β-hexachlorocyclohexane, and hexa-chlorobenzene | Decrease MDI (24 months) $PON1_{108TT}$ –5.7 (–9.0 to –2.5) 9=0.01; Decrease PDI (24 months) $PON1_{108TT}$ –2.8 (–5.7 to 0.2) p=0.07; increased odds PDD 2.0 (0.8 to 5.1) p=0.14; no association $PON1_{192}$; no association PON1 activity measured newborn, 2 years, maternal and MDI, PDI, PDD. Evidence of decreasing MDI score by number of $PON1_{108}$ variant alleles: $PON1_{108CC}$ –3.2 (–9.8 to 3.5), CT –3.7 (–8.0 to 0.6), TT –5.5 (–11.1 to 0.1), p-interaction 0.98. | In this study population, evidence PON1 may influence MDI score, but not PDI or PDD risk at two-years. Non-significant evidence of decreasing MDI score by increasing DAP levels across strata of the number of $PON1_{108}$ variant alleles, interaction non-significant. Similar trend with prenatal DEP levels and MDI, PDI by PON108 alleles, but less pronounced. Overall, limited, non-definitive evidence of effect modification by PON1 status in the relation between mental and psychomotor effects and prenatal DAPs. | Strengths: Longitudinal design, measurement and consideration of many confounders (including other environmental chemicals); the prenatal exposure measures were, with some exceptions, the average of two measurements, and thus may better reflect chronic exposure during the pregnancy<br><br>Weaknesses: Potential for residual confounding, potential for exposure misclassification, potential selection bias (not all children had the outcome measures), study population not generalizable to the whole US. Study may be under-powered to evaluate effect modification by *PON1* status. |
| Article 12: Marks *et al*. (2010) CHAMACOS (N=348) | CBCL; K-CPT; ADHD confidence index; Hillside behavioral rating scale; | DAP (geometric mean) pregnancy 109.0 nmol/L; DEP 17.7 nmol/L | Psychometrician, exact age at assessment, sex, maternal education, depressive symptoms, PPVT (continuous), ≥ | Prenatal DAPs were non-significantly associated with maternal report of attention problems and | Marked effect modification by gender: 11-fold increase ADHD composite indicator in boys, less than 2-fold in | Strengths: Longitudinal design, measurement and consideration of many confounders (including |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | composite ADHD indicator | | 15 hr out-of-home child care/week, breast feeding duration (months), maternal age, parity, marital status, active/passive smoking exposure and regular alcohol use during pregnancy, presence of father in home, maternal work status, and household income | ADHD at age 3.5 years, but were significantly related at age 5 years [CBCL attention problems: β = 0.7 points; 95% confidence interval (CI), 0.2-1.2; ADHD: β = 1.3; 95% CI, 0.4-2.1]. | girls, however unstable estimates; weak evidence of association DAPs at 3.5, 5 years and attention | other environmental chemicals); the prenatal exposure measures were, with some exceptions, the average of two measurements, and thus may better reflect chronic exposure during the pregnancy<br><br>Weaknesses: Potential for residual confounding, potential for exposure misclassification, potential selection bias (not all children had the outcome measures), study population not generalizable to the whole US. |
| Article 13: Rauh *et al*. (2011) Columbia U. (N=265) | • Wechsler Scales of Intelligence for Children (WISC-IV)<br>• Child Behavior Checklist (CBCL). | • Chlorpyrifos levels in umbilical cord blood samples, N=256 newborns<br>• If no cord blood (12% of subjects), levels were imputed from mothers' values.<br>• Values for samples with non-detectable chlorpyrifos levels (N=115, 43%) were imputed by using assay-specific limit of detection (LOD) values to impute an approximate level. | Data were collected regarding lead exposure, demographics, education and occupational history, income, active and passive smoking, alcohol and drug use during pregnancy, and residential pesticide use. Final models included prenatal environmental tobacco smoke (ETS) exposure, gender, ethnicity, gestational age at birth, quality of home care-taking environment, maternal education, and maternal IQ. | Full-Scale IQ: (B) of -0.003, CI = 0.006, 0.001, p= 0.064 Working Memory Index: (B) of -0.006, CI = 0.009, 0.002, p<0.001. The investigators articulated these results as showing that a 1 pg/g increase in chlorpyrifos exposure was associated with a 0.006 point decrease in the log-transformed Working Memory score and a 0.003 point decrease in the log-transformed Full-Scale IQ score. The investigators concluded that for each standard deviation increase in exposure (4.61pg/g) there is a 1.4% reduction in Full-Scale IQ and a 2.8% reduction in Working Memory. | For each standard deviation increase in exposure (4.61pg/g) there is a 1.4% reduction in Full-Scale IQ and a 2.8% reduction in Working Memory. | Strengths<br>• Direct assessment of chlorpyrifos levels using maternal serum and cord blood.<br>• Analysis using a continuous CPF level, which, in contrast to dichotomous CPF levels, provides a more meaningful look at potential threshold effects and dose-response trends.<br>• The investigators rigorously evaluated their methods for imputing values for undetectable CPF levels which in the end, were validated.<br>• The authors describe an elegant and methodologically sound statistical analysis, |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | addressing many of the potential shortcomings of their exposure data and covariates.<br><br>Weaknesses: The use of a single snapshot of prenatal chlorpyrifos exposure may not be an accurate surrogate for full prenatal exposure levels.<br>• There is no control for exposure over the subsequent 7 years which may be critical, especially as the process of neurocognitive development is fluid and rapid during these early childhood years. • Possibility of that an increased awareness of the risks of pesticide exposures could disproportionately affect postnatal exposure behavior.• Complicating this analysis is the pervasive, non-specific nature of neurological effects and the difficulty in attributing causal pathways. • when closely reviewed, the 95% CI for Full Scale IQ for both techniques contain 0 (LASSO: -0.006, 0.001, p=0.064; fully-adjusted: -0.006, 0.001, p=0.048• The authors do not address the clinical relevance of the 1.4% and 2.8% reductions and how this may impact a child or his/her |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | psychological or educational plans. |
| Article 9: Engel *et al*. (2011)<br>Mt. Sinai (N=169) | Wechsler Preschool and Primary Scale of Intelligence, 3rd edition (WPPSI-III) at ages < 7 years; Wechsler-IV Intelligence Scale (verbal comprehension; perceptual reasoning, working memory, processing speed, full scale intelligence) at age 7-9 years | DEP: 24.7 nmol/L; Total DAP: 82 nmol/L (same as Engel 2007) | Maternal age, race/ethnicity, marital status, education, breast-feeding, child sex, alcohol, smoking, or drug use during pregnancy, maternal IQ, a score based on assessment of the home environment (HOME), season of urine collection, language spoken in the home, age at testing, examiner and urinary creatinine level. | At age 6-9 years, non-statistically significant reductions in full scale IQ, perceptual reasoning, verbal comprehension, working memory and processing speed with increasing DAP, more profound with DEP than DMP | Used non-specific biomarker DEP/DAP; some evidence of effect modification by PON1 genotype | Limitations include use of non-specific markers of chlorpyrifos pesticide exposure (DAPs), use of only a single (third-trimester) urine sample, and the large proportion of loss to follow-up.<br><br>The statistical analysis was largely appropriate. Imputing of missing data may affect precision of association estimates and result in attenuated effect estimates as a result of exposure measurement error, although these are offset by the further categorization of the exposure levels (at the median).<br><br>Assessment and control for confounding were appropriate. However, confounding by unmeasured (and mis-measured) risk factors for abnormal growth that are related to pesticide exposure would bias the results in this study. Such factors may be related to socioeconomic status of the study participants, a cofactor which is difficult to define, no less measure in an epidemiologic study. |
| Article 14: | Wechsler-IV | Total DAPs | Maternal intelligence, measures | The authors observed | Prenatal measures taken | Strengths: the longitudinal |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| Bouchard *et al*. (2011) CHAMACOS (N=329) | Intelligence Scale (verbal comprehension; perceptual reasoning, working memory, processing speed, full scale intelligence) measured at age 7 years | (quintiles): Q1 (39 nmol/L); Q2 75 nmol/L; Q3 126 nmol/L; Q4 221 nmol/L; Q5 508 nmol/L. Geometric mean DAP 131 nmol/L | of how stimulating the environment is, and known or suspected neurotoxins were measured prenatally. Maternal intelligence was assessed via the Peabody Picture Vocabulary Test (PPVT). To measure the quality and extent of stimulation available to a child in the home environment, the Infant-Toddler HOME (Home Observation for Measurement of the Environment) inventory was completed at the 6-month, 1, 2, 3.5, 5, and 7 year visits; known or suspected neurotoxicants, polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs), p,p'-dichlorodiphenyltrichlorethane (DDT), p,p'-dichlorodiphenyltrichlorethylene (DDE), and lead. | evidence of an association between prenatal exposures to OP pesticides as measured by urinary DAP metabolites in women during pregnancy, and decreased cognitive functioning in children at age 7. | later half of pregnancy more significantly associated intelligence than early; little evidence post-natal OP exposure associated with intelligence; 7 point reduction in full scale intelligence DAP Q5/Q1 (SS) | design, the measurement of urinary DAP at multiple time points and following children to age seven when tests of cognitive function are reportedly more reliable. The authors were able to adjust for or consider many factors related to cognitive function, such as prenatal exposure to other environmental agents, socioeconomic indicators, maternal intelligence and education, and child stimulation. The cohort had a relatively homogenous socioeconomic profile, reducing the potential for uncontrolled confounding. |
| Article 15: Whyatt *et al*. (2007) Columbia U. (N=102) | None | Geometric mean, 6.9 ± 17.0 ng/m$^3$; range < 0.4–171 ng/m$^3$. Personal air monitor: median 2.8 ng/m$^3$, mean 6.2 ± 11.1 ng/m$^3$, range < 0.4–83.4 ng/m$^3$ | N/A | There was little within-home variability and no significant difference in air concentrations within homes over time ($p \geq 0.2$); between-home variability accounted for 88% of the variance in the indoor air levels of propoxur, 92% in chlorpyrifos, 94% in diazinon, and 62% in piperonyl butoxide ($p < 0.001$). Indoor and maternal personal air insecticide levels were highly correlated ($r = 0.7–0.9$, $p < 0.001$). | Indoor and maternal personal air insecticide levels were highly correlated ($r = 0.7–0.9$, $p < 0.001$). | Strengths: study design and exposure assessment techniques, Limitations: only those cohort participants enrolled after 2011 were included in the analysis (most likely due to the lack of serial data from the earlier years.) |
| Article 16: Whyatt *et al*. (2009) | None | The limit of detection (LOD) | N/A | Meconium TCPy concentrations were | TCPy in maternal urine samples was not reliable, | Comprehensive exposure assessment including actual |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| Columbia U. (N=102) | | of chlorpyrifos in blood samples was 0.5–1 pg/g plasma. The LOD of TCPy in urine samples was 0.26 ng/mL urine. The LOD for TCPy in meconium was 0.2 ng based on a sample weighing 0.5 g. Exposure marker levels below the LOD were given a value of half the level of detection, and were then log10 transformed. | | positively correlated with chlorpyrifos in maternal and cord blood ($r = 0.25–0.33$, $p < 0.05$) and with TCPy in maternal urine ($r = 0.31$, $p < 0.01$). | but the maternal and cord blood chlorpyrifos as well as the TCPy levels in meconium were reliable measures of exposure | blood chlorpyrifos levels, the repeated sampling, and the environmental sampling.<br><br>Weaknesses: only included participants recruited in the post-cancellation period, use of nonparametric, rank-based statistics is appropriate but the large number of observations below the level of detection receiving equal rank, may be problematic; no dietary assessment |
| Rauh et al.(2012), (n=40) | Morphological change in the pediatric brain in regions of the brain known to be associated with learning, cognition and social behavior | Tertile 3 ( ≥4.39 pg/g), compared to Tertiles 0, 1, 2 (<4.39 pg/g, including those not exposed to CPF) | Age, sex | Authors report differences in brain structure (regional cerebral size and thickness) by CPF exposure groups, and the differences (high>low CPF) in regional brain size is likely due to enlargement of underlying white matter. Statistical interaction by gender reported. | Authors concluded that the evidence from the study illustrated changes in brain morphology in association with higher CPF exposure, and that changes observed were in areas of the brain that subserve those learning, cognition and social behavioral, supported by previous observational and experimental literature. | Study supports general hypothesis of CPF influence on brain morphology, but lacks specific hypotheses regarding particular areas of the cerebrum affected; limited and somewhat unbalanced depiction of the available rodent experimental data; statistical methods appropriate, correction for multiple statistical comparisons a strength; MRI image readers blinded to exposure status enhances study validity; lack of information on other validation practices; small sample size, pilot study, low statistical power; external validity limited; one time measure |

| Author, Year, Sample Size | Outcome Assessed | Exposure Measurement | Potential confounders considered | Primary Result | Conclusions/Uncertainties | Strengths and Weaknesses |
|---|---|---|---|---|---|---|
| | | | | | | of pre-natal exposure |

### 2.2.10 Summary & Discussion

Across the eight studies identified in this 2015 literature review within the scope of this analysis and with adequate quality in design and exposure assessment, it is important to note that these studies used study methods that were highly variable, including different exposure measurement, outcome assessment, study design, and geographical location. These differences make it challenging to compare the results across studies. In comparison to the studies from the three US birth cohorts, results from the new studies (e.g., ELEMENT, CHARGE) provide some supportive evidence for the findings from CCCEH, Mt. Sinai and CHAMACOS but in general the new studies are not as robust as those from CCCEH, Mt. Sinai and CHAMACOS.

In the two Chinese studies, EPA does not know how the OP exposures from these studies relate to the currently registered use pattern for OPs used in the U.S./North America. In the case of these two Chinese studies (Guodong *et al.,* 2012; Zhang *et al.*, 2014), there may be differences in the study population and outcome measurements that may account for the observed differences in study results, with Zhang *et al*. (2014) documenting statistically significant associations for total DEAPs, total DMAPs, and total DAPs and Guodong *et al.* (2012) observing no association with these exposures. The Zhang *et al.* (2014) study was conducted in Shenyang, with a study population reported as 87% urban and 13% rural, whereas the Guodong *et al.* (2012) study was conducted in Shanghai with a 99% urban and 1% suburban study population. Given the higher percentage of study participants from rural areas, the study participants from Zhang *et al.* (2014) may have had different pattern and magnitude of OP exposures compared to those from Guodong *et al.* (2012). In addition, it is noted that different outcome measurements were made in these studies, with Guodong *et al.* (2012) assessing 23-25 month old children using a developmental quotient score and Zhang *et al.* (2014) assessing 3 day old infants using a Neonatal Behavioral Neurological Assessment. Given the different outcome assessments, exposure potential, study designs (cohort vs. cross-sectional), and ages of the participants in these two studies, it is difficult to draw conclusions on how these study results compare. It is notable that the results from the Zhang *et al.* (2014) study focusing on neonates are consistent with those from other studies which reported statistically significant associations between delayed neurological development measured in newborns and total DEAP, total DMAP, and total DAP exposure (Engel *et al.*, 2007; Young *et al.*, 2005). However, it is noted that neurological development was measured within a few days of birth for Zhang *et al.* (2014) and Engel *et al.* (2007), whereas Young *et al.* (2005) conducted their measurements within two months of birth.

In this 2015 literature review, two studies were identified that documented either suggestive or statistically significant associations between OP exposure and autism spectrum disorders (Furlong *et al.*, 2014; Shelton *et al.*, 2014). Specifically, Furlong *et al.* (2014) reported suggestive, but not statistically significant, evidence of an association between total DEAP exposure and reciprocal social responsiveness among black participants and boys. These results are consistent with previous studies conducted on the CCCEH and CHAMACOS cohorts, with these studies also showing statistically significant associations between OP exposure and

ASD (Rauh *et al.*, 2006; Eskenazi *et al.*, 2007).[7] Specifically, Eskenazi *et al.* (2007) reported a statistically significant association between PDD and total DAP exposure, whereas Rauh *et al.* (2006) showed a significant association between PDD and specifically chlorpyrifos exposure. Both PDD and reciprocal social responsiveness are related to the autism spectrum disorder. Using a different exposure assessment method (geospatial analysis and residential proximity to total OP exposure), Shelton *et al.* (2014) also documented statistically significant associations between total OP exposure and ASD. While these studies vary in the magnitude of the overall strength of association, they have consistently observed a positive association between OP exposure and ASD.

A total of three studies focusing on ADHD/behavioral/attention problems were identified in this 2015 literature review (Oulhote and Bouchard, 2013; Bouchard *et al.*, 2010; Fortenberry *et al.*, 2014). It is noted that the Fortenberry *et al.* (2014) study is a prospective cohort study on Mexican children, with this study adding an additional North American cohort to the previously identified U.S. cohort studies (CHAMACOS, Mt. Sinai, and CCEH). Of the four studies focusing on ADHD and OP exposure, three found statistically significant associations, with only Oulhote and Bouchard (2013) finding no association with total DMAP, DEAP, or total DAP exposure. In contrast, Bouchard *et al.* (2010) observed an association with total DMAP and total DAP exposure and ADHD. Fortenberry *et al.* (2014) found suggestive, but not statistically significant, evidence of an association with TCPy and ADHD in boys. Overall, the Fortenberry and Bouchard study results are consistent with that of earlier studies from the CCCEH (Rauh *et al.*, 2006) and CHAMACOS (Eskenazi *et al.*, 2007; Marks *et al.*, 2010). Specifically, statistically significant associations were observed by Rauh *et al.* (2006) with chlorpyrifos exposure and ADHD, Eskenazi *et al.* (2007) with total DMAPs and total DAPs and ADHD, and Marks *et al.* (2010) with total DEAP, DMAP, and total DAP exposure.

It is important to put into context the specific outcome measures used in the assessment of attention and neurobehavioral problems. For example, Bouchard *et al.* (2010) identified statistically significant associations between OP exposure and ADHD/behavioral problems, whereas Oulhote and Bouchard (2013) did not. It is valuable to compare these studies given that they are both cross-sectional studies using large population level datasets with biomarker information, with Oulhote and Bouchard (2013) using a Canadian dataset and Bouchard et al (2010) using a U.S. dataset. Bouchard *et al.* (2010) used criteria for ADHD from DSM-IV, whereas Oulhote and Bouchard (2013) used a "Strengths and Differences Questionnaire (SDQ)," with the SDQ being a more generic assessment of mental health status than the DSM-IV criteria. When Oulhote and Bouchard (2013) compared their results to Bouchard *et al.* (2010), they noted that their outcome measurements may not have been as sensitive and that this may account for the difference in study results.

[7] The DSM-V defines ASD (autism spectrum disorder) which now encompasses several disorders that were different diagnoses in DSM-IV, including PDD (pervasive developmental disorder, a catch-all where the other categories didn't fit). Depending on when the study was conducted, the authors may use the PDD or ASD criteria and terminology.

Across epidemiology studies looking at ADHD/behavioral problems, a suggestive or statistically significant positive association was observed in multiple studies between OP exposure and these neurobehavioral outcomes (Bouchard *et al.*, 2010; Fortenberry *et al.*, 2014; Rauh *et al.*, 2006, Eskenazi *et al.*, 2007). While these studies have differences in the years that the exposure occurred, study design, exposure assessment, and outcome assessment, the commonality in their results is striking.

When all the evidence is considered together, there are uncertainties with respect to a number of factors such as exposure assessment, lack of ability to make strong causal linkages, and unknown window(s) of susceptibility. The epidemiology studies reviewed in the 2012/2014 and 2015 literature reviews represent different investigators, locations, points in time, exposure assessment procedures, and outcome measurements. Despite all these uncertainties and differences in study design, authors consistently identified associations with neurodevelopmental outcomes such as ADHD/behavioral problems and autism spectrum, in relation to OP exposure.

## 3.0 Weight of Evidence Analysis: Integration Across Multiple Lines of Evidence

In 2010, OPP developed a draft "Framework for Incorporating Human Epidemiologic and Incident Data in Health Risk Assessment" which provides the foundation for evaluating multiple lines of scientific evidence in the context of the understanding of the adverse outcome pathway (or mode of action (U.S. EPA, 2010). The draft framework, which includes two key components: problem formulation and use of the MOA/AOP frameworks, was reviewed favorably by the SAP in 2010 (FIFRA SAP, 2010) and has recently been applied to chlorpyrifos. This document extends the chlorpyrifos WOE to other OPs.

One of the key components of the agency's draft framework is the use the MOA/AOP concept as a tool for organizing and integrating information from different sources to inform the causal nature of links observed in both experimental and observational studies. Specifically, the modified Bradford Hill Criteria (Hill, 1965) are used to evaluate the experimental support that establishes key events within MOA/AOP, and explicitly considers such concepts as strength, consistency, dose response, temporal concordance and biological plausibility in a weight of evidence analysis; sections 3.1-3.3 below summarize the available evidence based on these principles.

### 3.1 Dose-Response Relationships & Temporal Concordance

Since the MOA(s)/AOP(s) is/are not established for neurodevelopmental outcomes (USEPA, 2012, 2014), it is not possible to describe the concordance in key events or biological steps leading to neurodevelopmental outcomes. As such, the quantitative linkages between molecular initiating events (MIE)s, intermediate steps, and ultimately the adverse outcome (i.e., neurodevelopmental effects) cannot be determined.

With respect to the timing of exposure, across the epidemiology database of studies the maternal urine, cord blood and other (meconium) measures provide evidence that exposure did occur to the fetus during gestation, but the actual level of such exposure during the critical window(s) of susceptibility is not known. While significant uncertainties remain about the actual exposure levels experienced by mothers and infant participants in the children's health cohorts, it is unlikely that these exposures resulted in AChE inhibition. As part of the CHAMACOS study, Eskenazi *et al*. (2004) measured AChE activity and showed that no differences in AChE activity were observed. The biomarker data from the Columbia University studies are supported by the agency's dose reconstruction analysis using the PBPK-PD model. Following the recommendation of the FIFRA SAP (2012), the agency conducted a dose reconstruction analysis of residential uses available prior to 2000 for pregnant women and young children inside the home (USEPA, 2014). Based on the output from the PBPK-PD model, for the highest exposure considered (i.e., indoor broadcast use of a 1% chlorpyrifos formulation), <1% RBC AChE inhibition in pregnant women would be expected. While uncertainty exists as to actual OP exposure at (unknown) critical windows of exposure, EPA believes it is unlikely individuals in the epidemiology studies experienced RBC AChE inhibition.

Within the Columbia University epidemiology studies, the relationship in time between prenatal chlorpyrifos exposure and adverse neurodevelopmental outcomes is concordant. The time period under study within the Columbia University (CCCEH) study, spanned the point in time in which pesticide manufacturers voluntarily cancelled the use of chlorpyrifos in the home environment, and researchers were able to show the change in exposure before (high use period) and after (low/no use period) the period of removal of chlorpyrifos products from the residential marketplace. Moreover, prior to the voluntary cancellation there were >80% detectable levels of chlorpyrifos in cord blood but in the time period after the cancellation only 16% of the measured values were greater than the level of detection (LOD); there was only one child born in the time period subsequent to the voluntary cancellation of chlorpyrifos in the residential marketplace for whom the cord blood chlorpyrifos level was in the upper-tertile of pre-cancellation exposure levels. The significantly reduced proportion of measured values greater than the LOD as well as the observation of an absence of an association between prenatal chlorpyrifos exposure among infants born after the voluntary cancellation of chlorpyrifos and neurodevelopmental effects support the hypothesis that chlorpyrifos is related to these outcomes. However, as noted by study authors, EPA and the FIFRA SAP (2012), this could also be due to inadequate sample size to detect a small to modest effect among the group of infants born after the voluntary cancellation. It is notable that epidemiology studies from other research groups have not included analyses across different years of exposure.

### 3.2 Strength, Consistency & Specificity

Published and submitted laboratory animal studies have been reviewed for OPs. The >30 papers on chlorpyrifos provide evidence of long-lasting neurodevelopmental disorders in rats and mice; however, there was no clear consistency in terms of pattern, timing, or dose

response for these effects. The additional toxicological literature and guideline DNT studies with the other OPs provide more evidence for the same conclusions, with again the same caveats and uncertainties. While overall cognitive function and motor activity appeared to be altered the most often, it is apparent that these behaviors were also the most often evaluated.

Among the epidemiology studies, two of the cohorts (CCCEH and ELEMENT) have focused on chlorpyrifos whereas the other studies (Mt. Sinai cohort, CHAMACOS cohort, CHARGE study, Bouchard *et al*., 2007) have focused on less specific biomarkers (i.e., DAPs) and are not specific to any particular OP. When considered in concert, the epidemiology studies provide consistent findings for some outcomes. Specifically, with regard to the three US children's environmental health epidemiology studies and the ELEMENT cohort in Mexico, each of the four study cohorts utilized a prospective birth cohort study design in which mothers were recruited into study prior to the birth of the infants and development and identification of adverse effects. As noted above, the CHAMACOS and Mt. Sinai cohorts that measured neurological effects at birth (the Brazelton index), observed a putative association with OPs (Engel *et al*., 2007; Young *et al*., 2005). Similarly, while not consistent by age at time of testing (ranging from 6 month to 36 months across the three cohorts), the three US cohorts each reported evidence of impaired mental and psychomotor development. Attentional problems and ADHD were reported by CCCEH, Mt. Sinai, CHAMACOS, and ELEMENT investigators with additional support from Bouchard *et al*. (2010). In addition, several studies have now documented suggestive or positive associations between OP exposure and autism spectrum disorders (Rauh *et al*., 2006; Shelton *et al*., 2014; Furlong *et al.*, 2014; Eskenazi *et al.*, 2007; Eskenazi et al, 2010). Finally, each of the three US childrens cohort study authors observed an inverse relation between the respective prenatal measures of chlorpyrifos and intelligence measures at age 7 years.

As stated in the EPA neurotoxicity guidelines[8], direct extrapolation of developmental neurotoxicity results from laboratory animals to humans is limited by the lack of knowledge about underlying toxicological mechanisms and the relevance of these results to humans. EPA notes consistencies across the databases of *in vivo* laboratory animal studies and epidemiology studies, although challenges of making a direct comparison between neurodevelopmental domain inter-species remain. It can be assumed that developmental neurotoxicity effects in animal studies indicate the potential for altered neurobehavioral development in humans, although the specific types of developmental effects seen in experimental animal studies may not be the same as those that may be produced in humans. However, considering the toxicological and epidemiological data in the context of three major neurodevelopmental domains (specifically, cognition, motor control, and social behavior), insights can be gained. Previously reviewed studies of chlorpyrifos in rats and/or mice reported impaired cognition, changes in locomotor activity levels, altered social interaction, and to a lesser extent, changes in neuromotor function (FIFRA SAP 2012; USEPA, 2014). While there are fewer studies for all the other OPs, behavioral effects in the same functional domains were again reported. The most commonly reported outcome was cognitive dysfunction, and although it was overall consistent there were again differences in cognitive specificity, gender differences, or dose

[8] http://www.epa.gov/raf/publications/pdfs/NEUROTOX.PDF

response. Quite a few studies also report changes in motor activity and sensory function in offspring, but there generally fewer studies that assess social interactions for OPs other than chlorpyrifos. It is notable that the laboratory animal studies vary in experimental designs such as species, strain, gender, dosing regimens (age, routes, vehicle), and test parameters (age, protocol). Likewise, observational epidemiology studies vary by population characteristics (race/ethnicity, SES, and pesticide use/exposure profile), co-exposures (mix of chemicals, windows of exposure), and method of exposure and outcome assessment. Given the differences across laboratory animal and epidemiology studies, the qualitative similarity in research findings is striking.

In contrast, quantitatively, there are notable differences between animals and humans. Specifically, in animals, the doses most often used in these studies are sufficient to elicit ≥10% brain and RBC inhibition depending on the study design, age of the animal, and sampling time. In the epidemiology studies, based on the comparisons with biomonitoring data, reported AChE data from CHAMACOS and the results of the chlorpyrifos dose-reconstruction analysis, it is unlikely that RBC AChE would have been inhibited by any meaningful or measurable amount, if any at all, and most likely none in the brain. This key difference in dose response between the experimental toxicology and epidemiology studies poses challenges in interpreting such data. There are a number of possible hypotheses such as: 1) limitations of experimental laboratory studies which have limited statistical power due to relatively small sample sizes; 2) humans display a broader array of behaviors and cognitive abilities than rats, thus limiting the sensitivity of the rat studies; and 3) in the epidemiology studies, the timing of OP application and blood collections are not coupled—thus higher levels of blood OPs were likely missed.

### 3.3 Biological Plausability & Coherence

EPA's cancer guidelines (2005) includes guidance which are also applicable to this current evaluation of OPs. In fact, the Guidelines indicate:

"*evaluation of the biological plausibility of the associations observed in epidemiologic studies reflects consideration of both exposure-related factors and toxicological evidence relevant to identification of potential modes of action (MOAs). Similarly, consideration of the coherence of health effects associations reported in the epidemiologic literature reflects broad consideration of information pertaining to the nature of the biological markers evaluated in toxicologic and epidemiologic studies*. [p. 39]."

The Cancer Guidelines further state that "*lack of mechanistic data, however, is not a reason to reject causality* [p. 41]."

At this time, a MOA(s)/AOP(s) has/have not been established for neurodevelopmental outcomes. This growing body of literature does demonstrate, however, that OPs are biologically active on a number of processes that affect the developing brain. Moreover, there

is a large body of *in vivo* laboratory studies which show long-term behavioral effects from early life exposure, albeit at doses which cause AChE inhibition. EPA considers the results of the toxicological studies relevant to the human population, as qualitatively supported by the results of epidemiology studies. The lack of established MOA/AOP pathway does not undermine or reduce the confidence in the findings of the epidemiology studies. When all the evidence is considered together, there are uncertainties with respect to a number of factors such as exposure assessment, lack of ability to make strong causal linkages, and unknown window(s) of susceptibility. The epidemiology studies reviewed in the 2012/2014 and 2015 literature reviews represent different investigators, locations, points in time, exposure assessment procedures, and outcome measurements. Despite all these uncertainties and differences in study design, multiple investigators have identified associations with neurodevelopmental outcomes such as ADHD/behavioral problems and autism spectrum, in relation to OP exposure. There is evidence of delays in mental development in infants (24-36 months), attention problems and autism spectrum disorder in early childhood, and intelligence decrements in school age children who were exposed to OPs during gestation. Investigators reported strong measures of statistical association across several of these evaluations (odds ratios 2-4 fold increased in some instances), and observed evidence of exposures-response trends in some instances, *e.g.*, intelligence measures. Thus, with respect to biological plausibility and coherence, although uncertainties remain, these uncertainties are diminished in the context of the qualitative similarity between the epidemiology studies.

## 4.0 10X FQPA Safety Factor for Infants and Children

As section 408(b)(2)(C) of the FFDCA instructs EPA, in making its "reasonable certainty of no harm" finding, that in "the case of threshold effects, an additional tenfold margin of safety for the pesticide chemical residue and other sources of exposure shall be applied for infants and children to take into account potential pre- and postnatal toxicity and completeness of data with respect to exposure and toxicity to infants and children." Section 408 (b)(2)(C) further states that "the Administrator may use a different margin of safety for the pesticide chemical residue only if, on the basis of reliable data, such margin will be safe for infants and children." Given the totality of the evidence, there is sufficient uncertainty in the human dose-response relationship for neurodevelopmental effects which prevents the agency from reducing or removing the statutory 10X FQPA Safety Factor. For the preliminary human health risk assessments for the OPs, a value of 10X will be applied. Similarly, a database uncertainty factor of 10X will be retained for occupational risk assessments. The agency will continue to evaluate the epidemiology studies and pursue approaches for quantitative or semi-quantitative comparisons between doses which elicit AChE inhibition and those which are associated with neurodevelopmental outcomes prior to a revised human health risk assessment. **The FQPA 10X Safety Factor will be retained for OPs for the population subgroups that include infants, children, youths, and women of childbearing age for all exposure scenarios.**

## 5.0 References


Acosta-Maldonado, Brenda, Sánchez-Ramírez, Blanca, Reza-López, Sandra, and Levario-Carrillo, Margarita. Effects of exposure to pesticides during pregnancy on placental maturity and weight of newborns: A cross-sectional pilot study in women from the Chihuahua State, Mexico. Human and Experimental Toxicology 28(8) 451–459, 2009.

al-Hachim, G.M., Fink, G.B., 1968a. Effect of DDT or parathion on open-field behavior of offspring from DDT- or parathion-treated mothers. Psychol Rep **22**, 1193-1196.

al-Hachim, G.M., Fink, G.B., 1968b. Effect of DDT or parathion on condition avoidance response of offspring from DDT or parathion-treated mothers. Psychopharmacologia **12**, 424-427.

Andersen, Helle R., Debes, Fróði, Wohlfahrt-Veje, Christine, Murata, Katsuyuki, Grandjean, Philippe. Occupational pesticide exposure in early pregnancy associated with sex-specific neurobehavioral deficits in the children at school age. Neurotoxicology and Teratology 47:1-9, 2015.

Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., *et al*. (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem, 29*(3), 730-741.

Barr DB, Ananth CV, Yan X, Lashley S, Smulian JC, Ledoux TA, Hore P, Robson MG. Pesticide concentrations in maternal and umbilical cord sera and their relation to birth outcomes in a population of pregnant women and newborns in New Jersey. Sci Total Environ. 2010;408:790–795.

Berkowitz GS, Wetmur JG, Birman-Deych E, Obel J, Lapinski RH, Godbold JH, Holzman IR, Wolff MS. In Utero Pesticide Exposure, Maternal Paraoxonase Activity, and Head Circumference. Environmental Health Perspectives. 2004;112:388–391.

Boobis, A. R., Cohen, S. M., Dellarco, V., McGregor, D., Meek, M. E., Vickers, C., *et al*. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit Rev Toxicol, 36*(10), 781-792.

Boobis, A. R., Doe, J. E., Heinrich-Hirsch, B., Meek, M. E., Munn, S., Ruchirawat, M., *et al*. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit Rev Toxicol, 38*(2), 87-96.

Bouchard M.F., Bellinger D.C., Wright R.O., Weisskopf M.G., (2010). Attention-Deficit/Hyperactivity Disorder and Urinary Metabolites of Organophosphate Pesticides. PEDIATRICSDOI:10.1542/peds.2009-3058.

Bouchard, M. F., Chevrier, J., Harley, K. G., Kogut, K., Vedar, M., Calderon, N., *et al*. (2011). Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. Environ Health Perspect, 119(8), 1189-1195.

Ceh, K., Jezek, D., Zuzek, M.C., Cebulj-Kadunc, N., Majdic, G., 2012. Effects of organophosphate chlormephos on development and function of testes and brain. Slovenian Veterinary Research **49**, 41-49.

Chen L, Zhao T, Pan C, Ross JH, Krieger RI. (2012) Preformed biomarkers including dialkylphosphates (DAPs) in produce may confound biomonitoring in pesticide exposure and risk assessment. J Agric Food Chem. 2012 Sep 12;60(36):9342-51. doi: 10.1021/jf303116p. Epub 2012 Sep 4.

Clemens, G.R., Hartnagel, R.E., Bare, J.J., Thyssen, J.H., 1990. Teratological, neurochemical, and postnatal neurobehavioral assessment of METASYSTOX-R, an organophosphate pesticide in the rat. Fundam Appl Toxicol **14**, 131-143.

Crowder, L.A., Lanzaro, G.C., Whitson, R.S., 1980. Behavioral effects of methyl parathion and toxaphene exposure in rats. J Environ Sci Health B **15**, 365-378.

Dąbrowski, Sławomir, Hanke, Wojciech, Polańska, Kinga, Makowiec-Dąbrowska, Teresa, and Sobala, Wojciech. Pesticide exposure and birthweight: An epidemiological study in Central Poland. International Journal of Occupational Medicine and Environmental Health 16(1): 31-39, 2003.

de Castro, V.L., Chiorato, S.H., Pinto, N.F., 2001. Relevance of developmental testing of exposure to methamidophos during gestation to its toxicology evaluation. Toxicol Lett **118**, 93-102.

Engel SM, Berkowitz GS, Barr DB, Teitelbaum SL, Siskind J, Meisel SJ, Wetmur JG, Wolff MS. Prenatal Organophosphate Metabolite and Organochlorine Levels and Performance on the Brazelton Neonatal Behavioral Assessment Scale in a Multiethnic Pregnancy Cohort. American Journal of Epidemiology. 2007;165:1397-1404.

Engel SM, Wetmur J, Chen J, Zhu C, Barr DB, Canfield RL, Wolff MS. Prenatal Exposure to Organophosphates, Paraoxonase 1, and Cognitive Development in Childhood. Environmental Health Perspectives. 2011;119:1182-1188.

Eskenazi B, Harley K, Bradman A, Weltzien E, Jewell NP, Barr DB, Furlong CE, Holland NT. Association of in Utero Organophosphate Pesticide Exposure and Fetal Growth and Length of Gestation in an Agricultural Population. Environmental Health Perspectives. 2004;115:792–798.

Eskenazi B, Marks AR, Bradman A, Harley K, Barr DB, Johnson C, Morga N, Jewell NP. Organophosphate Pesticide Exposure and Neurodevelopment in Young Mexican-American Children. Environmental Health Perspectives. 2007;115:792–798.

Eskenazi, B., Huen, K., Marks, A., Harley, K. G., Bradman, A., Barr, D. B., & Holland, N. (2010). PON1 and neurodevelopment in children from the CHAMACOS study exposed to organophosphate pesticides in utero. Environ Health Perspect, 118(12), 1775-1781.

FIFRA Scientific Advisory Panel. (2008b). "The Agency's Evaluation of the Toxicity Profile of Chlorpyrifos." Report from the FIFRA Scientific Advisory Panel Meeting of September, 2008. FIFRA Scientific Advisory Panel, Office of Science Coordination and Policy, Office of Prevention, Pesticides and Toxic Substances, U.S. Environmental Protection Agency. Washington, DC. Available: http://www.epa.gov/scipoly/sap/meetings/2008/index.html

FIFRA Scientific Advisory Panel. (2010). February 2 - 4, 2010: Incorporation of Epidemiology and Human Incident Data into Human Risk Assessment.

FIFRA Scientific Advisory Panel. (2012). Transmittal of the Meeting Minutes from the FIFRA Scientific Advisory Panel Meeting of April 10-12, 2012. FIFRA Scientific Advisory Panel, Office of Science Coordination and Policy, Office of Prevention, Pesticides and Toxic Substances, U.S. Environmental Protection Agency. Washington, DC. Available: http://www.epa.gov/scipoly/sap/meetings/2012/041012meeting.html

Fortenberry G.Z., Meeker J.D., Sanchez B.N., *et al*., (2014). Urinary 3,5,6-tichloro-2-pyridinol (TCPY) in pregnant women from Mexico City: Distribution, temporal variability, and relationship with child attention and hyperactivity. International Journal of Hygiene and Environmental Health 217 (2014) 405-412.

Fortenberry, GZ, Meeker, JD, Sánchez, BN, Barr, DB, Peterson, K, Bellinger, D, Schnaas, L, Solano-González, M, Ettinger, AS, Hernandez-Avila, M, Hu, H, Tellez-Rojoj, MM. Paraoxonase I polymorphisms and attention/hyperactivity in school-age children from Mexico City, Mexico. (2014a.) Environmental Research 132:342–349.

Furlong, Melissa A., Engel, Stephanie M., Boyd Barr, Dana, Wolff, Mary S. Prenatal exposure to organophosphate pesticides and reciprocal social behavior in childhood. 2014. Environment International 70:125-131.

Grandjean, Philippe, Harari, Raul, Barr, Dana B., Debes, Frodi. Pesticide exposure and stunting as independent predictors of neurobehavioral deficits in Ecuadorian school children. 2006. Pediatrics 117(3):e546-e556.

Guodong D., Pei W., Ying T., Jun Z., *et al*., (2012). Organophosphate Pesticide Exposure and Neurodevelopment in Young Shanghai Children. Environ Sci. Technol. 2012, 46, 2911-2917.

Gupta, R.C., Rech, R.H., Lovell, K.L., Welsch, F., Thornburg, J.E., 1985. Brain cholinergic, behavioral, and morphological development in rats exposed in utero to methylparathion. Toxicol Appl Pharmacol **77**, 405-413.

Handal A.J., Lozoff B., Breilh J., Harlow S.D., (2007). Effect of community of residence on neurobehavioral development in infants and young children in a flower-growing region of Ecuador. Environ Health Perspect. 2007 Jan;115(1): 128-33.

Handal A.J., Lozoff B., Breilh J., Harlow S.D., (2007B). Neurobehavioral Development in Children with Potential Exposure to Pesticides. Epidemiology May 2007; 18: 312-320.

Handal A.J., Lozoff B., Breilh J., Harlow S.D., (2008). Occupational Exposure to Pesticides During Pregnancy and Neurobehavioral Development of Infants and Toddlers. Epidemiology, Volume 19, number 6, November 2008.

Harari, Raul, Julvez, Jordi, Murata, Katsuyuki, Barr, Dana, Bellinger, David C., Debes, Frodi, and Grandjean, Philippe. Neurobehavioral Deficits and Increased Blood Pressure in School-Age Children Prenatally Exposed to Pesticides. 2010. Environmental Health Perspectives 118(6): 890-896.

Harley KG, Huen K, Schall RA, Holland NT, Bradman A, *et al*. Association of Organophosphate Pesticide Exposure and Paraoxonase with Birth Outcome in Mexican-American Women. PLoS ONE 2011; 6(8)1-10.

Hill AB (1965). The Environment and Disease: Association or Causation? Proc R Soc Med. May 1965; 58(5): 295–300.

Johnson, F.O., Chambers, J.E., Nail, C.A., Givaruangsawat, S., Carr, R.L., 2009. Developmental chlorpyrifos and methyl parathion exposure alters radial-arm maze performance in juvenile and adult rats. Toxicol Sci **109**, 132-142.

Kofman, Ora, Berger, Andrea, Massarwa, Ali, Friedman, Alon, and Abu Jaffar, Abed. Motor Inhibition and Learning Impairments in School-Aged Children Following Exposure to Organophosphate Pesticides in Infancy. Pediatric Research. 60:88-92, 2006.

Koutroulakis, D., Sifakis, S., Tzatzarakis, M.N., Alegakis, A.K., Theodoropoulou, E., Kavvalakis, M.P., Kappou, D., Tsatsakis, A.M. Dialkyl phosphates in amniotic fluid as a biomarker of fetal exposure to organophosphates in Crete, Greece; association with fetal growth. Reproductive Toxicology 46 (2014) 98–105.

Lazarini, C.A., Lima, R.Y., Guedes, A.P., Bernardi, M.M., 2004. Prenatal exposure to dichlorvos: physical and behavioral effects on rat offspring. Neurotoxicol Teratol **26**, 607-614.

Lehotzky, K., Szeberenyi, M.J., Kiss, A., 1989. Behavioral consequences of prenatal exposure to the organophosphate insecticide sumithion. Neurotoxicol Teratol **11**, 321-324.

Levin, E.D., Timofeeva, O.A., Yang, L., Petro, A., Ryde, I.T., Wrench, N., Seidler, F.J., Slotkin, T.A., 2009. Early postnatal parathion exposure in rats causes sex-selective cognitive impairment and neurotransmitter defects which emerge in aging. Behav Brain Res **208**, 319-327.

Lima, C.S., Dutra-Tavares, A.C., Nunes, F., Nunes-Freitas, A.L., Ribeiro-Carvalho, A., Filgueiras, C.C., Manhaes, A.C., Meyer, A., Abreu-Villaca, Y., 2013. Methamidophos exposure during the early postnatal period of mice: immediate and late-emergent effects on the cholinergic and serotonergic systems and behavior. Toxicol Sci **134**, 125-139.

Lizardi, Patricia Sanchez, O'Rourke, Mary Kay, and Morris, Richard J. The effects of organophosphate pesticide exposure on Hispanic children's cognitive and behavioral functioning. Journal of Pediatric Psychology 33(1): 91–101, 2008.

Llop S., Julvez J., Marina L.S., Vizcaino E., *et al.*, (2013). Prenatal and postnatal insecticide use and infant neuropsychological development in a multicenter birth cohort study. Environment International 59(2013) 175-182.

Lovasi GS, Quinn JW, Rauh VA, Perera FP, Andrews HF, Garfinkel R, Hoepner L, Whyatt R, Rundle A. Chlorpyrifos Exposure and Urban Residential Environment Characteristics as Determinants of Early Childhood Neurodevelopment. American Journal of Public Health. 2010;101(1):63-70. doi: 10.2105/AJPH.2009.168419. Epub 2010 Mar 18.

Lu, *et al*. Assessing the association between pesticide exposure and cognitive development in rural Costa Rican children living in organic and conventional coffee farms. Int J Adolesc Med Health. 2009; 21(4): 609-621

Marks AR, Harley K, Bradman A, Kogut K, Barr DB, Johnson C, *et al*. Organophosphate Pesticide Exposure and Attention in Young Mexican-American Children: The CHAMACOS Study. Environmental Health Perspectives. 2010;118:1768-1774.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J, Vickers C. 2014. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J Appl Toxicol.* 2014 Jan;34(1):1-18.

Moreno-Banda, G., Blanco-Muñoz, J., Lacasaña, M., Rothenberg, S.J., Aguilar-Garduño, C., Gamboa, R., Pérez-Méndez, O. Maternal exposure to floricultural work during pregnancy, PON1 Q192R polymorphisms and the risk of low birth weight. Science of the Total Environment 407: 5478–5485, 2009.

Muñoz-Quezada MT, Lucero BA, Barr DB, Steenland K, Levy K, Ryan PB, Veronica Iglesiase V, Alvaradoe S, Conchaf C, Rojas E, Catalina Vega C. Neurodevelopmental effects in children associated with exposure to organophosphate pesticides: A systematic review. Neurotoxicology 39: 158-168.

Naksen W, Prapamontol T, Mangklabruks A, Chantara S, Thavornyutikarn P, Srinual N, Panuwet P, Ryan PB, Riederer AM, Barr DB. Associations of maternal organophosphate pesticide exposure and PON1 activity with birth outcomes in SAWASDEE birth cohort, Thailand. Environmental Research142: 288–296, 2015.

Nevison. A comparison of temporal trends in United States autism prevalence to trends in suspected environmental factors. Environmental Health. 2014; 13(73): 16 pp

Oulhote, Youssef and Bouchard, Maryse F. Urinary Metabolites of Organophosphate and Pyrethroid Pesticides and Behavioral Problems in Canadian Children. Environmental Health Perspectives. 121:1378-1384 (2013)

Petit, Claire, Chevrier, Cécile, Durand, Gaël, Monfort, Christine, Rouget, Florence, Garlantezec, Ronan, Cordier, Sylvaine. Impact on fetal growth of prenatal exposure

to pesticides due to agricultural activities: a prospective cohort study in Brittany, France. Environmental Health 9:71-83, 2010.

Quiros-Alcala L., Alkom A.D., Boyce W.T., Lippert S., *et al*., (2011). Maternal prenatal and child organophosphate pesticide exposures and children's autonomic function. NeuroToxicology 32 (2011) 646-655.

Racke J.D. Degradation of organophosphorous insecticidesin environmental matrices. In: Organophosphates Chemistry, Fate and Effects. Academic Press, Inc., San Diego, CA, 1992, pp 47–72.

Rauch, Stephen A., Braun, Joe M., Boyd Barr, Dana, Calafat, Antonia M., Khoury, Jane, M., Montesano, Angela, Yolton, Kimberly, and Lanphear, Bruce P. Associations of Prenatal Exposure to Organophosphate Pesticide Metabolites with Gestational Age and Birth Weight. Environmental Health Perspectives. 120:1055–1060 (2012).

Rauh VA, Garfinkel R, Perera FP, Andrews HF, Hoepner L, Barr DB, Whitehead R, Tang D, Whyatt RW. Impact of Prenatal Chlorpyrifos Exposure on Neurodevelopment in the First 3 Years of Life Among Inner-City Children. Pediatrics. 2006;18(6): 1835-1859.

Rauh V, Arunajadai S, Horton M, Perera F, Hoepner L, Barr DB, *et al*. 2011. Seven-Year Neurodevelopmental Scores and Prenatal Exposure to Chlorpyrifos, a Common Agricultural Pesticide. Environmental Health Perspectives. 2011;119:1196-1201.

Rauh, V. A., Perera, F. P., Horton, M. K., Whyatt, R. M., Bansal, R., Hao, X., . . . Peterson, B. S. (2012). Brain anomalies in children exposed prenatally to a common organophosphate pesticide. Proc Natl Acad Sci U S A, 109(20), 7871-7876.

Roegge, C.S., Timofeeva, O.A., Seidler, F.J., Slotkin, T.A., Levin, E.D., 2007. Developmental diazinon neurotoxicity in rats: later effects on emotional response. Brain Res Bull **75**, 166-172.

Rohlman, Diane S., Arcury, Thomas A., Quandt, Sara A., Lasarev, Michael, Rothlein, Joan, Travers, Rachelle, Tamulinas, Alys, Scherer, Jennifer, Early, Julie, Marín, Antonio, Phillips, Jacki, McCauley, Linda. Neurobehavioral performance in preschool children from agricultural and non-agricultural communities in Oregon and North Carolina. Neurotoxicology 26:895-598, 2005.

Ruckart, *et al*. Long-Term Neurobehavioral Health Effects of Methyl Parathion Exposure in Children in Mississippi and Ohio. Environmental Health Perspectives. 2004; 112(1):46-51.

Samarawickrema, Nirma, Pathmeswaran, Arunasalam, Wickremasinghe, Rajitha, Peiris-John, Roshini, Karunaratna, Madawa, Buckley, Nick, Dawson, Andrew, and De Silva, Janaka. Fetal effects of environmental exposure of pregnant women to organophosphorus compounds in a rural farming community in Sri Lanka. Clinical Toxicology 46:489–495, 2008.

Savitz, David A., Arbuckle, Tye, Kaczor, Diane, and Curtis, Kathryn M. Male pesticide

exposure and pregnancy outcome. American Journal of Epidemiology 146(12):1025-1036, 1997.

Shelton, Janie F., Geraghty, Estella M., Tancredi, Daniel J., Delwiche, Lora D., Schmidt, Rebecca J., Ritz, Beate, Hansen, Robin L., and Hertz-Picciotto, Irva. Neurodevelopmental disorders and prenatal residential proximity to agricultural pesticides: The CHARGE Study. Environmental Health Perspectives 122:1103–1109, 2014.

Simon TW, Simons SS Jr, Preston RJ, Boobis AR, Cohen SM, Doerrer NG, Fenner-Crisp PA, McMullin TS, McQueen CA, Rowlands JC; RISK21 Dose-Response Subteam. 2014. The use of mode of action information in risk assessment: Quantitative key events/dose-response framework for modeling the dose-response for key events. *Crit Rev Toxicol.* 2014 Aug;44 Suppl 3:17-43.

Slotkin, T.A., Tate, C.A., Ryde, I.T., Levin, E.D., Seidler, F.J. 2006. Organophosphate insecticides target the serotonergic system in developing rat brain regions: Disparate effects of diazinon and parathion at doses spanning the threshold for cholinesterase inhibition. Environ Health Perspec **114**, 1542-1546.

Sochaski, M.A., McManus, B.M., Struve, M.F., Wallace, D.G., Dorman, D.C. 2007. Inhibition and recovery of maternal and foetal cholinesterase enzymes following fenitrothion administration in CD rats. Xenobiotica **37**, 19-29.

Spyker, J.M., Avery, D.L., 1977. Neurobehavioral effects of prenatal exposure to the organophosphate Diazinon in mice. J Toxicol Environ Health **3**, 989-1002.

Stamper, C.R., Balduini, W., Murphy, S.D., Costa, L.G., 1988. Behavioral and biochemical effects of postnatal parathion exposure in the rat. Neurotoxicol Teratol **10**, 261-266.

Suarez-Lopez J. R., Jacobs D.R., Himes J.H., Alexander B.H., Lazovich D., *et al*., (2012). Lower acetylcholinesterase activity among children living with flower plantation workers. Environmental Research 114 (2012) 53-59.

Suarez-Lopez J. R., Jacobs D.R., Himes J.H., Alexander B.H., (2013). Aectylcholinesterase activity and neurodevelopment in boys and girls. PEDIATRICS Volume 132, number 6, December 2013.

Suarez-Lopez J. R., Jacobs D.R., Himes J.H., Alexander B.H., (2013a). Acetycholinesterase activity, cohabitation with floricultural workers, and blood pressure in Ecuadorian Children. Environmental Health Perspectives, volume 121, number 5, May 2013.

Timofeeva, O.A., Roegge, C.S., Seidler, F.J., Slotkin, T.A., Levin, E.D., 2008a. Persistent cognitive alterations in rats after early postnatal exposure to low doses of the organophosphate pesticide, diazinon. Neurotoxicol Teratol **30**, 38-45.

Timofeeva, O.A., Sanders, D., Seemann, K., Yang, L., Hermanson, D., Regenbogen, S., Agoos, S., Kallepalli, A., Rastogi, A., Braddy, D., Wells, C., Perraut, C., Seidler, F.J., Slotkin, T.A., Levin, E.D., 2008b. Persistent behavioral alterations in rats neonatally

exposed to low doses of the organophosphate pesticide, parathion. Brain Res Bull **77**, 404-411.

U.S. Environmental Protection Agency. (2005). Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum, U.S. Environmental Protection Agency. Washington, DC. EPA/630/P-03/001F.

U.S. Environmental Protection Agency. (2006). Revised Organophosphorous Pesticide Cumulative Risk Assessment; July 31, 2006. Office of Pesticide Programs, U.S. Environmental Protection Agency. Washington, D.C. Available http://www.epa.gov/pesticides/cumulative/rra-op.

U.S. Environmental Protection Agency. (2008). Draft Science Issue Paper: Chlorpyrifos Hazard and Dose Response Characterization. http://www.epa.gov/scipoly/sap/meetings/2008/091608_mtg.htm

U.S. Environmental Protection Agency. (2010). Draft Framework for Incorporating Human Epidemiologic and Incident Data in Health Risk Assessment, January 7, 2010.

U.S. Environmental Protection Agency. (2011a). Chlorpyrifos: Preliminary Human Health Risk Assessment for Registration Review. June 30, 2011. D388070.

U.S. Environmental Protection Agency. (2011b). Chlorpyrifos: Occupational and Residential Exposure Assessment, EPA Barcode: D388165,. In Office of Pesticide Programs (Ed.). Washington D.C.

U.S. Environmental Protection Agency. (2012). Draft Issue Paper: Scientific Issues Concerning Health Effects of Chlorpyrifos. http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2012-0040-0002

U.S. Environmental Protection Agency. (2014). Chlorpyrifos: Revised Human Health Risk Assessment for Registration Review. December 29, 2014. D424485.

Vatanparast, J., Naseh, M., Baniasadi, M., Haghdoost-Yazdi, H., 2012. Developmental exposure to chlorpyrifos and diazinon differentially affect passive avoidance performance and nitric oxide synthase-containing neurons in the basolateral complex of the amygdala. Brain Res 1494, 17-27.

Wang, Pei, Tian, Ying, Wang, Xiao-Jin, Gao, Yu, Shi, Rong, Wang, Guo-Quan, Hu, Guo-Hua, Shen, Xiao-Ming. Organophosphate pesticide exposure and perinatal outcomes in Shanghai, China. 4 Environment International 2:100–104, 2012.

Whyatt, R. M., Barr, D. B., Camann, D. E., Kinney, P. L., Barr, J. R., Andrews, H. F., . . . Perera, F. P. (2003). Contemporary-use pesticides in personal air samples during pregnancy and blood samples at delivery among urban minority mothers and newborns. Environ Health Perspect, 111(5), 749-756.

Whyatt RM, Rauh V, Barr DB, Camann DE, Andrews HF, Garfinkel R, Hoepner LA, Diaz D, Dietrich J, Reyes A, Tang D, Kinney PL, Perera FP. Prenatal Insecticide Exposures and Birth Weight and Length among an Urban Minority Cohort. Environmental

Health Perspectives. 2004;112:1125–1132.

Whyatt RM, Garfinkel R, Hoepner LA, Holmes D, Borjas M, Williams MK, *et al*. Within- and Between-Home Variability in Indoor-Air Insecticide Levels during Pregnancy among an Inner-City Cohort from New York City. Environmental Health Perspectives 2007;115:383-389.

Whyatt RM, Garfinkel R, Hoepner LA, Andrews H, Holmes D, Williams MK, *et al*. A Biomarker Validation Study of Prenatal Chlorpyrifos Exposure within an Inner-City Cohort during Pregnancy. Environmental Health Perspectives. 2009;117:559-567.

Wickerham, Erin L., Lozoff, Betsy, Shao, Jie, Kaciroti, Niko, Xia, Yankai, Meeker, John D. Reduced birth weight in relation to pesticide mixtures detected in cord blood of full-term infants. Environment International 47:80–85, 2012.

Win-Shwe, T.T., Nakajima, D., Ahmed, S., Fujimaki, H., 2012. Impairment of novel object recognition in adulthood after neonatal exposure to diazinon. Arch Toxicol **87**, 753-762.

Wolff, MS, Engel, S, Berkowitz, G, Teitelbaum, S, Siskind, J, Barr, DB, Wetmur, J. Prenatal Pesticide and PCB Exposures and Birth Outcomes. 2007. Pediatric Research 61(2):243-250.

Young JG, Eskenazi B, Gladstone EA, Bradman A, Pedersen L, Johnson C, Barr DB, Furlong CE, Holland NT. Association Between In Utero Organophosphate Pesticide Exposure and Abnormal Reflexes in Neonates. NeuroToxicology. 2005;26:199–209.

Zhang Y., Han S., Liang D., Shi X., Wang F., *et al*., (2014). Prenatal exposure to organophosphate pesticides and neurobehavioral development of neonates: A birth cohort study in Shenyang, China. Plos ONE 9(2):e88491. doi:10.1371/journal.pone.0088491.

Zhang X, Driver JH, Li Y, Ross JH, Krieger RI. (2008) Dialkylphosphates (DAPs) in fruits and vegetables may confound biomonitoring in organophosphorus insecticide exposure and risk assessment. J Agric Food Chem. 2008 Nov 26;56(22):10638-45. doi: 10.1021/jf8018084.

## 6.0 Appendices

1. Table of *In Vivo* Developmental Neurotoxicity Studies of OPs.
2. Summary of Guideline DNT Studies Submitted to the Agency for OPs other than Chlorpyrifos.
3. Low Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment

## Appendix 1. Table of *In Vivo* Developmental Neurotoxicity Studies of OPs.

(Effects described are only those measured after weaning. Bold indicates functional domains that were reported to show treatment effects.)

| OP | Study | Species & strain | Dose, route, vehicle | Dosing period | ChE inhibition | Domain | Age of testing | Outcomes | NOEL, LOEL | Notes & Study Problems |
|---|---|---|---|---|---|---|---|---|---|---|
| Chlormephos | Ceh *et al*., 2012 | mouse BALB/c | 3.5, 0.35 ug/ml in drinking water of dams ~ 0.6, 0.06 mg/kg/d (@ 5ml/d, 30 g) | 7 day premating to weaning | No | **Anxiety & Emotion** | PND70-80 | Increased time in closed arms & decreased time in open arms in elevated plus maze, ~0.6 mg/kg/d, M&F | NOEL=0.35 ug/ml in water, ~0.06 mg/kg/d | Litter not unit of statistical analysis<br>No pup allocation described but had to have used some littermates<br>M & F responses appear similar but not statistically compared |
| Diazinon | Roegge *et al*., 2008 | rat Sprague Dawley | 0.5, 2 mg/kg/d sc DMSO | PND1-4 | Cites Slotkin *et al*. (2006): 0.5 mg/kg/d produced <10% brain inhibition on PND5, 2 mg/kg/d produced 25-30% brain inhibition 2 hr after dose on PND4, and 10-20% inhibition on PND5 | **Anxiety & Emotion** | PND52-56 | Decreased time in open arms in elevated plus maze, 2 mg/kg/d, M only | No NOEL LOEL=0.5 mg/kg/d | Pups & dams redistributed daily<br>No effect in F<br>Accepts p<0.1 as significant for interactions<br>No dose-response in feeding or milk preference studies<br>Abstract misstates sex differences |
| | | | | | | | PND64-67, 78-79 | Decreased latency to eat in novelty suppressed but not home-cage feeding, 0.5 & 2 mg/kg/d, M only | | |
| | | | | | | | PND73-74 | Decreased chocolate milk preference, 0.5 mg/kg/d only, M only | | |
| | | | | | | | PND86-87 | No effect on forced swim test | | |
| Diazinon | Spyker and Avery, 1977 | mouse F2 hybrid (NCTR cross bred) | 0.18, 9 mg/kg/d in peanut butter | GD1-birth | No | **Sensory** | PND38 | Increased errors on visual cliff, 0.18 mg/kg/d only, F only<br>No effect acoustic startle or olfactory responses | No NOEL LOEL=0.18 mg/kg/d | No pup allocation described but had to have used littermates<br>Not clear when both sexes tested and/or compared<br>Statistics not described<br>Maternal weight gain lowered at both doses<br>Weight gain of high dose pups decreased<br>Preweaning testing: decreased contact placing, 0.18 mg/kg/d only<br>No dose-response for some measures<br>Twice as many controls as treated<br>Looks like decreased rotarod endurance PND65, not significant due to high variability |
| | | | | | | **Neuromotor** | PND50, 60, 65,75 | No effect swimming ability<br>Increased rod cling endurance, 0.18 & 9 mg/kg/d, sex not specified<br>Decreased inclined plane performance, 0.18 & 9 mg/kg/d, sex not specified | | |
| | | | | | | Activity | PND75-76 | No effect open field | | |
| | | | | | | Cognition | PND87 | No effect errors in Lashley maze | | |
| Diazinon | Timofeeva *et al*., 2008 | rat Sprague Dawley | 0.5, 2 mg/kg/d sc DMSO | PND1-4 | Cites Slotkin *et al*. (2006): 0.5 mg/kg/d produced <10% brain inhibition on PND5, 2 mg/kg/d produced | Activity | PND28-42 | No effect in figure-8 chamber, M&F | No NOEL LOEL=0.5 mg/kg/d | Littermates of those used in Roegge *et al.* 2008<br>Pups & dams redistributed daily<br>Accepts p<0.1 as significant for interactions |
| | | | | | | **Sensory** | PND77-84 | Decreased prepulse inhibition, 0.5 & 2 mg/kg/d, M only | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 25-30% brain inhibition 2 hr after dose on PND4, and 10-20% inhibition on PND5 | **Cognition** | PND28-35, 91-126 | No effect on T-maze spontaneous alternation, M&F<br>Increased working memory errors, 0.5 mg/kg/d only, M&F | | No mention of sex effects in T-maze<br>No dose-response for some measures |
| Diazinon | Vatanparast *et al.*, 2013 | rat Wistar | 1 mg/kg/d sc DMSO | GD15-18 or PND1-4 | No | Activity | PND60 | Gestational: No effect in open field, M&F<br>Postnatal: No effect in open field, M&F | No NOEL<br>LOEL=1 mg/kg/d | 1 M & 1 F per litter but sex not nested within litter in statistics<br>F only affected with gestational exposure, both sexes affected with postnatal, looks like M more affected<br>Large effect sizes<br>Number of dams not mentioned<br>Discrepancy in text on pup sample sizes |
| | | | | | | **Cognition** | PND60-63 | Gestational: Decreased latency to cross and increased time spent in dark side on retention trial, no effect acquisition in passive avoidance, F only<br>Postnatal: Decreased latency to cross and increased time spent in dark side on retention trial, no effect acquisition in passive avoidance, M&F | | |
| Diazinon | Win-Shwe *et al.*, 2013 | mouse C3H/HeN | 0.5, 5 mg/kg/d sc DMSO | PND8-11 | No | **Cognition** | PND46-49, 81-84 | PND46-49: Decreased novel object exploration and discrimination, 0.5 & 5 mg/kg/d<br>PND81-84: Decreased novel object exploration and discrimination, 5 mg/kg/d only | No NOEL<br>LOEL=0.5 mg/kg/d | M only tested<br>Separate mice at two test times<br>Litter allocation to dose group not described<br>Assumes that ChE inhibition reported by Slotkin in rats would be same as in mice<br>Larger sample size at later age |
| Dichlorvos | Lazarini *et al.*, 2004 | rat Wistar | 8 mg/kg po (dilution of technical product); vehicle from formulation | GD6-15 | No | **Activity** | PND21, "adult" | PND21: Decreased locomotion open field, M only<br>"Adult": Decreased locomotion and increased immobility, only M tested | No NOEL<br>LOEL=8 mg/kg/d | Data analyzed as litter but sex not nested within litter in statistics<br>No effect on physical and reflex preweaning development<br>Only M tested as after PND21<br>Adult age not given |
| | | | | | | **Cognition** | "Adult" | Decreased latency to cross on retention trial in passive avoidance | | |
| Fenitrothion (sumithionR, 50% ai) | Lehotzky *et al.*, 1989 | rat Lati | 5, 10, 15 mg/kg/d po sunflower oil | GD7-15 | No | **Neuromotor** | PND26, 36, 104 | Decreased latency to fall off rotarod PND26, 104, not 36, 15 mg/kg/d | NOEL=5 mg/kg/d<br>LOEL=10 mg/kg/d | Postnatal mortality at all doses (16-17.5%)<br>No pup allocation described but had to have used some littermates<br>Only M tested<br>Statistics not described<br>Measured startle, righting, contact placing on PND22 but no results given<br>Shorter latency in cognitive task hard to interpret<br>Nonsignificant decrease in activity at PND26 |
| | | | | | | **Activity** | PND26, 36, 104 | Decreased activity in open field PND104, 15 mg/kg/d | | |
| | | | | | | **Cognition** | PND42, 104 | Shorter escape latency in conditioned response during acquisition, 10 & 15 mg/kg/d | | |
| | | | | | | **Social behavior** | PND62 | Increased time in social interaction, 10 & 15 mg/kg/d | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Methamidophos | deCastro *et al.*, 2000 | rat Wistar | 1 mg/kg/d po water | GD6-15 | Pilot in nonpregnant F dosed for 10 d gives 17% plasma inhibition at 1 mg/kg/d | Activity | PND40 | No effect in open field, sex no mentioned | NOEL=1 mg/kg/d | Used 2 pups/litter but no mention of sex, apparently used as independent observations<br>No effect on preweaning swimming performance<br>Decreased immobility time PND 14 only<br>Open field measures with really high variability, not reliable |
| Methamidophos | Lima *et al.*, 2013 | mouse Swiss | 1 mg/kg/d sc DMSO | PND3-9 | Pilot showed for 1 mg/kg/d: ~19% brain inhibition on PND10; ~36%, 46% brain inhibition 1, 4 hr after dosing on PND3; ~53, 61% brain inhibition 1, 4 hr after dosing on PND9; no brain inhibition in PND60 adults | Activity | PND61 | No effect in open field, sex not mentioned | No NOEL LOEL=1 mg/kg/d | 1 M & 1 F per litter<br>No data for M & F separately or mention of statistical differences<br>Dosing by litter<br>High variability especially with passive avoidance |
| | | | | | | **Anxiety & Emotion** | PND60-61 | Increased immobility time in forced swim, sex not mentioned<br>No effect on elevated plus maze, sex not mentioned | | |
| | | | | | | Cognition | PND63 | No effect on passive avoidance, sex not mentioned | | |
| Methyl parathion | Crowder *et al.*, 1980 | rat Sprague Dawley | 1 mg/kg/d po corn oil | GD7-15 | No | **Activity** | PND23, 30, 44, 54, 65 | Increased activity in open field, only PND23 and 54, sex not mentioned<br>to criterion | No NOEL LOEL=1 mg/kg/d | Preweaning testing: possibly decreased wire cling time (not analyzed), no effect on righting, startle, placement response<br>Increased postnatal mortality (30%)<br>Littermates used<br>Only 3 litters used<br>Small sample size for maze testing<br>Statistics not mentioned except for maze transfer test, just used t-test<br>Sex not mentioned except for maze transfer test, data not given for M & F<br>Methods & results cursory |
| | | | | | | **Cognition** | >PND68 | Slower transfer on 1st, 4th direction change in T-maze learning transfer, sex not mentioned | | |
| Methyl parathion | Gupta *et al.*, 1985 | rat Wistar-Furth F mated with F344 M | 1 mg/kg/d in peanut butter, 1.5 mg/kg/d po peanut oil | GD6-20 | Dams on GD19 show 20, 60% brain inhibition at 1, 1.5 mg/kg/d<br>Pups show brain inhibition up to 50% on PND1, 7, 14, 21, 1 & 1.5 mg/kg/d; on PND28 only 1.5 mg/kg/d | **Cognition** | PND60 | No effect on passive avoidance<br>No effect on shuttle box avoidance<br>Slower latency to bar press & increased days to asymptote on operant task (no schedule given), 1 mg/kg/d only, sex not mentioned | No NOEL LOEL=1 mg/kg/d but no effects at 1.5 mg/kg/d | High dose dams had cholinergic signs, increased resorptions<br>Pups moved to foster mothers at 24 hr<br>No effect on preweaning reflexive behaviors<br>Pup allocation not described<br>Statistics barely described<br>M & F apparently tested but data for each not shown or mentioned<br>Methods cursory<br>No dose-response for behavior but there is dose-response for ChE inhibition |
| | | | | | | Neuromotor | PND60 | No effect on rotarod | | |
| | | | | | | **Activity** | PND60 | Decreased activity, 1 mg/kg/d only, sex not mentioned | | |
| | | | | | | **Anxiety & Emotion** | PND60 | Faster cage emergence, 1 mg/kg/d only, sex not mentioned | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Sensory | PND120 | No effect on acoustic startle response | | Only 4/dose for operant testing |
| Methyl parathion | Johnson *et al.*, 2009 | rat Sprague Dawley | Incrementing doses: low 0.2 mg/kg/d throughout; mid 0.2, 0.4, 0.6 mg/kg/d every 5-6 d; high 0.3, 0.6, 0.9 mg/kg/d every 5-6 d | PND1-21 | Low dose showed 13-15% brain inhibition and high dose showed 63, 20, 18% brain inhibition on PND20, 30, 40; all doses recovered by PND50 | **Cognition** | PND29-60 | Increased working memory errors, mid and high dose, M only<br>Increased reference memory errors, all doses, M only | No NOEL<br>LOEL=0.2 mg/kg/d | Split-litter dose design<br>Included litter as random effect in statistics<br>No effect on preweaning measures of reflex development<br>No effect in F |
| Oxydemeton methyl (metasystoxR, 91% ai) | Clemens *et al.*, 1990 | rat CD | 0.5, 1.5, 4.5 mg/kg/d water | GD6-15 | Dams on GD16 show 30, 54, 72% plasma inhibition (RBC similar) & 22, 52, 68% brain inhibition at 0.5, 1.5, 4.5 mg/kg/d<br>Dams on GD20 show 20, 39, 54% brain inhibition at 0.5, 1.5, 4.5 mg/kg/d, 40% RBC inhibition at 4.5 mg/kg/d, and no plasma inhibition<br>Fetuses on GD20 show no brain inhibition | Cognition | PND25, 26, 35 | No effect on M-maze | NOEL=4.5 mg/kg/d | High dose dams had tremors<br>1 M & 1 F per litter<br>No data for M & F separately or mention of statistical differences<br>Statistics barely described<br>No effect on preweaning reflex or sensory tests |
| | | | | | | Activity | | No effect in open field | | |
| Parathion | Al-Hachim & Fink, 1968 | mouse CF1 | 3 mg/kg/d po<br>Corn oil | 3 dosing times: 1st, 2nd, or 3rd trimester | No | Cognition | PND30-37 | No effect on conditioned avoidance learning | NOEL=3 mg/kg/d | Very similar experiment as other papers, maybe same study<br>Pup allocation not clear but littermates probably used<br>Inadequate statistics<br>No mention of sex |
| Parathion | Al-Hachim & Fink, 1968 | mouse CF1 | 3 mg/kg/d po<br>Corn oil | 3 dosing times: 1st, 2nd, or 3rd trimester | No | Activity | PND60-66 | No effect in open field | NOEL=3 mg/kg/d | Very similar experiment as other papers, maybe same study<br>Pup allocation not clear but littermates probably used<br>Inadequate stats<br>No mention of sex |
| Parathion | Levin *et al.*, 2010 | rat Sprague Dawley | 0.1, 0.2 mg/kg/d sc DMSO | PND1-4 | Cites Slotkin *et al.* 2006: 0.1 mg/kg/d produced 5-15% brain inhibition on PND5, no data for 0.2 mg/kg/d | **Cognition** | PND420, 510, 570 | PND420: Increased working memory errors in radial arm maze, 0.1 mg/kg/d, M only; increased reference memory errors, 0.1 & 0.2 mg/kg/d, M only<br>PND510: Increased working memory errors in radial arm maze, 0.1 & 0.2 mg/kg/d, M only<br>PND570: No effect in radial arm maze | No NOEL<br>LOEL=0.1 mg/kg/d | Littermates of those used in Timofeeva 2008<br>Pups & dams redistributed daily<br>5% mortality high dose<br>No effect in F<br>Accepts p<0.1 as significant for interactions<br>No dose-response for several measures |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parathion | Stamper *et al.*, 1988 | rat Long Evans | 1.3, 1.9 mg/kg/d sc corn oil | PND5-20 | 35, 68% brn inh PND21; 26, 36% brn in PND28 | Activity | PND24 | No effect in open field | No NOEL LOEL=1.3 mg/kg/d | Split-litter dose design<br>Pup allocation not clear but littermates probably used<br>High dose produced cholinergic signs, says doses are 33 and 50% of LD50 in PND5 rat<br>Decreased weight gain with both doses<br>Preweaning, increased cliff avoidance latency, no effect righting, negative geotaxis, open field<br>M only tested<br>No post-hoc comparison of groups when significant, but looks like effects in both doses<br>No dose-response in working memory errors |
| | | | | | | Neuromotor | PND24 | No effect on rotarod | | |
| | | | | | | **Cognition** | PND24, PND 35-37 | Decreased alternation rate in T-maze, 1.3 & 1.9 mg/kg/d, only M tested<br>Increased working memory errors, 1.3 & 1.9 mg/kg/d, only M tested | | |
| Parathion | Timofeeva *et al.*, 2008 | rat Sprague Dawley | 0.1. 0.2 mg/kg/d sc DMSO | PND1-4 | Cites Slotkin *et al.* 2006: 0.1 mg/kg/d produced 5-15% brain inhibition on PND5, no data for 0.2 mg/kg/d | Activity | PND58-61 | No effect in figure-8 chambers | No NOEL LOEL=0.1 mg/kg/d | Pups & dams redistributed daily<br>5% mortality high dose<br>Accepts p<0.1 as significant for interactions<br>All radial arm maze effects only in low dose |
| | | | | | | **Cognition** | PND35-45<br>PND112-182 | No effect in T-maze spontaneous alternation<br>Decreased working memory errors in radial arm maze, 0.1 mg/kg/d only, M&F | | |
| | | | | | | **Anxiety & Emotion** | PND50-53<br>PND64-72<br>PND81-94 | Increased time in open arms in elevated plus maze, 0.2 mg/kg/d, M&F<br>No effect on novelty suppressed feeding<br>No effect on chocolate milk preference | | |
| | | | | | | **Sensory** | PND78-81 | Decreased tactile startle, 0.2 mg/kg/d, M&F<br>No effect on prepulse inhibition | | |

## Appendix 2. Summary of Guideline DNT Studies Submitted to the Agency for OPs other than Chlorpyrifos.

(Only changes that were observed after exposure had ended (post weaning, adult) are listed. 'X' indicates no significant changes on tests for each domain.)

| Chemicals | Cognition | Motor activity | Acoustic startle | Neuromotor (FOB) | Notes |
|---|---|---|---|---|---|
| Acephate | X | X | X | X | |
| Azinphos-methyl | X | X | X | X | |
| Coumaphos | X | X | X | X | |

| | | | | | |
|---|---|---|---|---|---|
| Diazinon | Biel maze: increase errors & latency high dose (~33.1 mg/g,/d, dam diet) M, PND24 & PND62; also mid dose (~3.4 mg/kg/d, dam diet) F, PND24 | X | X | X | |
| Dichlorvos[1] | -- | -- | -- | -- | |
| Dicrotophos | X | X | X | X | |
| Dimethoate | X | X | X | X | |
| Disulfoton | X | X | X | X | |
| Ethoprop | M maze: increase trials to criterion high dose (~29.3 mg/kg/d, dam diet) M, PND60 | X | X | X | |
| Fenamiphos | X | X | X | X | |
| Malathion | X | increased rearing in FOB open field mid dose (50 mg/kg/d to dams & pups), F, significant at PND45 only (maybe also PND60); no change automated motor activity | increased peak amplitude later blocks (perhaps habituation effect) all doses (5, 50, 150 mg/kg/d to dams & pups), F only, PND23; increased peak amplitude without prepulse low dose only, F only, PND60 | altered gait mid & high dose (50, 150 mg/kg/d to dams & pups), M & F, PND60 but not earlier | |
| Methamidophos | ~X | X | decreased peak amplitude early blocks mid & high dose (~1.65, 5.2 mg/kg/d dam diet), F only, significant at PND38, looks same but not significant at PND60 | X | PA PND24: report says decreased latency but nothing significant & table is questionable; M maze PND60: report says increased trials to criterion (M) and increased (M) or decreased (F) errors, but nothing significant and table is questionable |
| Methyl parathion | X | X | X | X | |
| Naled | X | X | increased peak amplitude & decreased latency middle blocks low dose only (0.4 mg/kg/d to dams & pups), F only, PND60; report says decreased amplitude but not significant all blocks high dose (10 mg/kg/d to dams & pups), M only, PND23 & PND60 | X | only swimming time in Y maze reported, varied significances, no mention of errors or other performance measures |

| | | | | | |
|---|---|---|---|---|---|
| Phorate | M maze: decrease number reaching criterion with relearning, low & mid dose (0.03 & 0.1 mg/kg/d to dams & pups), M only significant, PND30; not seen in second study with higher dose | X | decreased peak amplitude all blocks high dose (0.1 mg/kg/d to dams & pups), M only, PND60; not seen in second study with higher dose | X | combined two studies; one with low & mid dose, other with high dose; some data didn't agree |
| Profenofos | X | X | X | X | |
| Terbufos | X | X | X | X | |
| Tetrachlorvinphos | X | X | X | X | |
| Tribufos | X | X | X | X | |
| Trichlorfon | PA: decreased latency to enter on retention high dose (~205.1 mg/kg/d dam diet), M only, PND29; M maze: increased average errors second trial high dose, F only, PND60 | decreased activity middle blocks (maybefaster habituation) mid dose (~76.2 mg/kg/d dam diet) only, F only, PND60 | decreased peak amplitude all blocks high dose (~205.1 mg/kg/d dam diet), early blocks mid dose (~76.2 mg/kg/d dam diet), M & F, PND22; decreased amplitude middle blocks (not consistent) all doses (~23.3, 76.2, 205.1 mg/kg/d dam diet), F only, PND38; for M decreased amplitude apparent but not significant high dose PND38 & PND60 | X | |

[1] high pup mortality in all groups, including control, negated any valid neurotoxicity assessments of dichlorvos

## Appendix 3. Low Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Acosta-Maldonado *et al*. (2009) | Chihuahua, Mexico | Singleton pregnancies | Cross-sectional, small pilot study – only 9 women exposed, participant selection and exclusion not detailed<br><br>N=54 mothers<br>(9 exposed, 45 comparison mothers) | Proxy indicator of exposure - Residence in agricultural community were pesticides had been applied; or home located < 5 km from a pesticide application zone; or cohabitating with worker exposed to pesticides or agricultural labor<br><br>Also, AChE activity – Objective biomarker of exposure/altered function | Standardized but partially subjective assessment of placental maturity | Minimal. Adjustment for placental characteristics. | Appropriate multivariate analysis; Corrected hypothesis test results for multiple comparisons. | Selection bias unlikely; Residual confounding likely; substantial potential for differential misclassification of exposure. |
| Dawbrowski *et al*. (2003) | Lodz, Poland | Newborn children among Polish farmers | Case-control, large sample size<br><br>N=389<br>Age: newborns | Prenatal pesticide exposure assessed via questionnaire (retrospective self-report). Site visit *after* delivery to evaluate pesticide exposure | Pregnancy outcomes assessed using birth records | Appropriate. Included maternal demographics, predictors of high-risk pregnancy (duration, maternal weight) and environmental toxicant exposure (ETS). No adjustment for SES indicators | Appropriate multivariate analysis. | Selection bias unlikely; Residual confounding likely small in magnitude; considerable potential for non-differential misclassification of exposure (unlikely to account for non-null findings). Minimal misclassification of outcome. |
| Grandjean *et al*.(2006) | Tabacundo, Ecuador | Healthy 2nd and 3rd grade children | Cross-sectional, small sample size<br><br>N=72<br><br>Age <9 years | Prenatal occupational exposure assessed via questionnaire; Also recent child exposure biomarker assessment (DAP). | Objective anthropometric and other clinical outcomes; Numerous neurobehavioral outcomes evaluated using easy-to-administer screening instruments; Age appropriate. May be insensitive to subtle effects of OP pesticide exposure effects. | Appropriate. Homogenous population limited confounding by design; Included child demographics (age, sex, weight); SES indicators (maternal race, housing, running water, sewage), diet (meals/day), environmental toxicants (maternal alcohol and smoking) and medical history. | Appropriate multivariate analysis; Numerous hypotheses evaluated without correction for multiple comparisons. | Selection bias unlikely; Residual confounding likely small in magnitude; considerable potential for non-differential misclassification of exposure (likely to account for null findings). Potential misclassification of outcome. |

**Appendix 3. Low Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Harari *et al.* (2010) | Tabacundo, Ecuador (Andean plateau north of Quito) | 6- to 8-year-old children in the two lowest grades (called second and third) of one of two schools | Cross-sectional, small sample size<br><br>n = 81<br>Age 6- 8 years | Child Current Exposure - DMP metabolites measured in a spot urine samples. A blood sample was analyzed for AChE.<br><br>Maternal Exposure – Interview by skilled interviewers | Blood pressure and neurophysiologic measures - the instruments used for neurophsychologic measures were validated to avoid cross-cultural influences. | Appropriate: child's sex, age, BMI, number of daily meals (only in current exposure), stunting, hematocrit, school grade, having repeated one grade, maternal education level, family living in a traditional house, drinking water supply, and paternal education and employment | Appropriate: Standard parametric tests and logistic regression | Errors in exposure classification status since it was based on the maternal self-report, mothers likely being aware of the neurobehavioral status of their children, exposure assessment based on a spot urine sample |
| Kofman *et al.* (2006) | Israel (Negev region) | Bedouin population (Children aged 6 to 12 at the time of the study, who were victims of poisoning before age three) | Retrospective cohort study, small sample size<br>N = 52<br>9-Exposed to OP; 17-Exposed to Kerosene/paint thinner<br>26-Controls<br>Aged 6-12 years | OP poisoning was confirmed by low serum butyrylcholinesterase activity based on hospital records | Neuropsychological evaluation and structured interview of parents. Errors in assessment minimized as psychologists were qualified individuals, language and cultural differences taken into consideration, each child tested on same day and in same place as matched control. | Age, sex, background (cultural and demographic) | Difference in means | Errors is outcome classification likely since psychologists who administered the tests knew which children were exposed, small sample size |
| Koutroulakis *et al.* (2014) | Crete, Greece | Women with singleton pregnancies, permanent residents for at least two years, referral for amniocentesis to the Fetal-Maternal Unit, Department of Obstetrics and Gynecology, University Hospital of Heraklion | Prospective Cohort Study – large sample size, ethical issues<br><br>n = 415<br>Age: newborns | Objective. DAP measurement in a single amniotic fluid samples collected at either 16th or 20th weeks of gestation – Novel biomarker. Questionnaire was also used. | Birth weight and head circumference. Unclear how the outcome information was obtained. | Neonatal sex, maternal age, agricultural activities, and gestational age at amniocentesis – Rationale for confounder selection not provided | Appropriate: multiple linear regression | Comparing exposure measurements in AF against known biomarker (e.g., OP metabolite levels in urine) for validation of AF not conducted. Smoking status of participants before/during pregnancy, high risk pregnancies, gestational diabetes, PON1 enzyme activity in the fetuses not considered. |

**Appendix 3. Low Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Lizardi *et al*. (2008) | Yuma County, Arizona, USA | Children's Pesticide Survey (CPS) | Cross-sectional, small sample size<br><br>N=48<br><br>Age 7 years | Objective Biomarker of prenatal OP pesticide exposure (DAP) quantified in single child spot urine sample provided at the time of the cognitive assessment. | Cognitive assessment using battery of test instruments considered valid and reliable in similar populations. Spanish translation as necessary. | None | Correlation coefficients (type unspecified). Statistically significant confounders were not robust due to influential outliers | Selection bias unlikely; Residual confounding likely; considerable potential for non-differential misclassification of exposure Potential misclassification of outcome. |
| Lu *et al*. (2009) | Cota Brus, Costa Rica | 4-10 yr. old children whose parents worked in organic coffee farm (La Amistad) and conventional coffee farms (Las Mellizas) | Cross-Sectional (pilot), low sample numbers N=35:<br>17 Organic farm<br>18 Conventional farms<br>Age 4-10 years old | Good measure (urinary PNP, IMPY, TCPy), but no major differences between exposure groups | Good measure (CBARS) but different SES and demographic characteristics for exposure groups | Limited number (group, age, sex, handedness, grade) | Appropriate: Linear mixed effects for significant test outcomes from paired t-test analyses | Somewhat high - Convenience sample with different recruitment methods; exposure misclassification |
| Moreno-Banda *et al*. (2009) | Mexico (Villa Guerrero, Coatepec de Harinas, Tenancingo (Mexico); Cuernavaca, Cuautla, Jiutepec, Temixco, (Morelos) | Newborn children of floricultural workers and families | Cross-sectional, large sample size<br><br>N=328<br><br>Age: newborns | Proxy indicator of prenatal occupational OP pesticide exposure (self-reported floricultural occupation) | Objectively measured birth outcomes assessed using birth certificate (partial). Self-reported by mother if birth certificate unavailable. | Appropriate, though minimal – history of adverse reproductive outcomes, infant sex, maternal smoking and alcohol use during pregnancy. | Appropriate multivariate analysis | Selection bias likely; Residual confounding likely small in magnitude; considerable potential for non-differential misclassification of exposure (likely accounts for null findings). Potential outcome misclassification (self-report in subset of participants). |
| Nevison (2014) | U.S. National population | Children with birth years 1970–2005 in 1) California Department of Developmental Services (CDDS) reports 2) Individuals with Disabilities Education Act (IDEA) reports | Ecological (time trend) N=NP (national database) | Non-specific, proxy OP exposure measure (lbs/yr) | Two large reporting DBs: CA Department of Developmental Services (CDDS), US Individuals with Disabilities Education Act (IDEA) | Limited: differences in autism definitions, changes in diagnostic criteria | Appropriate: ratio of age-resolved snapshot; tracking trend slopes; correlation coefficient between temporal trend and composite autism prevalence curve | Selection bias unlikely; do not know individual exposure |

**Appendix 3. Low Quality Studies: Summary of Study Design Elements Impacting Study Quality Assignment (continued)**

| Study | Study Location | Cohort Name / Description of Study Population | Study Design | Exposure Assessment | Outcome Assessment | Confounder / Covariate Control | Statistical Analysis | Risk of (other) Bias |
|---|---|---|---|---|---|---|---|---|
| Rohlman *et al.* (2005) | Oregon and North Carolina, USA | Latino children of immigrant parents living in Oregon or North Carolina | Cross-Sectional, small sample size<br><br>N= 78<br>Age 48-71 months | Proxy indicator of chronic OP pesticide exposure - Residence in highly agricultural communities | Battery of neurocognitive development; Screening tools; Some instruments likely not appropriate for use in study population - Not all participants able to complete all evaluations. Poor administration. Tests administered twice; only performance on 2nd evaluation considered in analysis. | Self-reported covariate information collected via questionnaire. Adjustment for age, SES indicator (maternal education). No environmental toxicants. | One-sided hypothesis tests. Numerous hypotheses evaluated without correction for multiple comparisons. | Selection bias unlikely; Residual confounding likely; substantial potential for differential misclassification of exposure and outcome. Inadequate presentation of study results. Seemingly post-hoc evaluation of effect modification. |
| Samarawickrema *et al.* (2008) | Southern Sri Lanka | Pregnant women delivering at Embilipitiya Base Hospital | Cross-sectional Birth Cohort, small sample size<br><br>N=41 end of two pesticide spraying seasons;<br><br>N=25 at beginning of spraying season | Proxy indicator of prenatal OP pesticide exposure (delivery during pesticide spray season); Objective pesticide biomarkers assessed (OP pesticide residues), but detected in only two subject's specimens – not evaluated | Objective biomarkers of early biological effect outcomes - Maternal and fetal butyrylcholinesterase (BuChE) activity; antioxidant status;<br><br>fetal oxidative stress;<br><br>fetal DNA fragmentation | No adjustment for potential confounders (though comparison groups were considered to be relatively homogeneous). | Largely appropriate. Assumptions of some statistical tests likely violated. | Selection bias unlikely; Residual confounding likely; substantial potential for differential misclassification of exposure; Seemingly post-hoc evaluation of effect modification. |
| Savitz *et al.* (1997) | Ontario, Canada | Newborn children in the Ontario Farm Family Health Study; | Retrospective Birth Cohort, large sample size<br><br>N= 1,898 couples; 3,984<br><br>Age: newborns | Proxy indicator of pre-conception paternal para-occupational OP pesticide exposure (self-reported male farm activities in 3-month period prior to conception. | Objectively measured birth outcomes assessed by maternal self-report | Appropriate. Included family/child demographics (sex, weight, maternal age, ethnicity); SES indicators (maternal and paternal education and occupation, per capita income) race, housing, running water, sewage), diet (meals/day), pregnancy risks (maternal caffeine, alcohol and smoking) and medical history. | Inappropriate multivariate regression analysis. Likely misspecification of true variance; No adjustment for multiple comparisons. | Selection bias probable; Residual confounding likely; substantial potential for differential misclassification of exposure and outcome. |
| Wickerham *et al.* (2012) | Zhejiang Province, China | Newborn children delivered at the Fuyang Maternal and Children's hospital | Cross-sectional, small pilot study<br><br>n=116<br><br>Age: Full term infants | Objective biomarker of pesticide exposure (pesticide residues in cord blood) – parameterized as number of pesticide residues detected. Methods unlikely suitable for detecting low levels. | Birth weight assessed using birth records and maternal report | Appropriate. Assessed using questionnaire (maternal self-report) and medical records. | Appropriate multivariate analysis. | Selection bias unlikely; Residual confounding likely; substantial potential for differential misclassification of exposure. Some outcome misclassification unlikely . |
| Naksen et al (2015) | Fang district, Chiang Mai province, Thailand | Pregnant women delivering at Fang Hospital | Prospective Cohort, small pilot study<br><br>n=52<br><br>Age: newborns | Objective biomarkers of pesticide exposure (DAPs). Also AChE, BChE, and PON1 genotype expression measurement. Maternal blood and urinary samples taken, plus cord blood. Questionnaire to assess other exposures and covariates. | Birth outcomes (Body weight and length, and head circumference) abstracted from medical records. | Appropriate. Assessed using questionnaire (maternal self-report) and medical records. | Some errors in statistical analysis were identified; e.g., for gestation age, log total DEAP at 32 weeks of pregnancy, a 0.7 beta was reported, but the confidence interval is reported as (-0.1, -1.4)]. No adjustment for multiple comparisons. | Inadequate presentation of study results. Selection bias possible due to loss to follow-up; Residual confounding likely, small in magnitude; potential for non-differential misclassification of exposure. |